

OM of: US-09-271-584A-2 to: EST:* out_format : pfs
 Date: Mar 26, 2001 9:51 AM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
 -MODEL=frame+p2n.model -DEV=xlp
 -O=/cgn2_1/USP70_spool/US09271584/runat_15032001_151301_5660/app_query.fasta_1.604
 -DB=EST -QFMT=fastap -SUFFIX=p2n.rst -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -DELEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=2000000000 -USER=US09271584_CGN1_1.2656 -NCPU=6
 -ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-271-584A-2

Query length: 538

Database: EST*

Database sequences: 7991742

Database length: -791223438

Search time (sec): 1231.570000

score_list:

Sequence	Strd Orig	zScore	EScore	Len	Documentation
gb_est71:BE420587	+ 1034.50	2078.33	9.4e-107	1212	! BE420587 HWM000.D12 ITEC HWM
gb_est80:CE9809	+ 1028.00	2072.77	1.9e-106	651	! C99909 C99909 Arabidopsis thali
gb_est48:AW685820	+ 714.50	1432.96	8.3e-71	653	! AW685820 NF035E04NR1F1000 Nodul
gb_est71:BE440835	+ 694.00	1394.38	1.2e-68	502	! BE440835 SP42F03.y1 Gm-cl043 G
gb_est73:BE438253	+ 684.00	1374.37	1.5e-67	486	! AI938253 SC41H02.y1 Gm-cl014 G
gb_est73:BE438253	+ 632.00	1268.69	1.2e-61	469	! BE438253 NXCI_048.H12.F NXCI (N
gb_est76:BE803390	+ 590.50	1181.50	8.5e-57	574	! BE803390 SF55G05.y1 Gm-cl051 G
gb_est36:AV408114	+ 582.00	1167.91	4.8e-56	424	! AV408114 AV408114 Lotus japonic
gb_est71:BE433982	+ 554.50	1109.28	8.9e-53	519	! BE433982 EST405060 tomato break
gb_est73:BE607262	+ 544.00	1091.05	9.2e-52	401	! BE607262 NXCI_034.G06.F NXCI (N
gb_est80:CE91832	+ 522.00	1046.88	2.7e-49	378	! CE91832 C91832 Rice panicle sho
gb_est48:AW694621	+ 496.50	986.24	6.4e-46	757	! AW694621 NF078C11ST1F1085 devel
gb_est70:BE403631	+ 492.50	980.71	1.3e-45	612	! BE403631 WHE0434.H04_0082S whea
gb_est73:BE600656	+ 463.00	921.70	2.5e-42	556	! BE600656 P11_91.A10.D1_A002 Pat
gb_est27:AI941126	+ 442.00	884.51	3.0e-40	352	! AI941126 SB84H02.y1 Gm-cl010 G
gb_est48:AW691011	+ 428.50	849.29	2.7e-38	654	! AW691011 NF040D08ST1F1000 Devel
gb_est86:T75860	+ 415.00	830.21	3.1e-37	330	! T75860 10638 Lambda-PRL2 Arabid
gb_est46:AW597260	+ 414.50	825.39	7.5e-37	527	! AW597260 S191C11.y1 Gm-cl031 G
gb_est46:AW562918	+ 395.50	785.32	9.9e-35	498	! AW562918 66009F09.xl 660 - Mix
gb_est27:AI933308	+ 381.00	749.28	1.0e-32	838	! AI933308 701496019 A. thaliana,
gb_est72:BE498756	+ 380.00	754.56	5.1e-33	464	! BE498756 WHE0965.C04_F072S Whea
gb_est74:BE704796	+ 378.50	746.41	1.5e-32	700	! BE704796 SC02_02F04.A SC02.AAFC
gb_est73:BE597622	+ 358.50	707.69	2.1e-30	591	! BE597622 P11_72.A11.D1_A002 Pat
gb_gss25:AG021043	+ 334.50	657.90	1.2e-27	631	! AG021043 Oryza sativa DNA, 3'
gb_est73:BE597324	+ 292.50	570.94	8.6e-23	698	! BE597324 P11_72.A11.g1_A002 Pat
gb_est36:AV383444	+ 289.50	569.17	1.1e-22	491	! AV383444 AV383444 Halocynthia r
gb_est69:BE321087	+ 285.00	553.00	8.6e-22	863	! BE321087 NF033H11RT1F1092 Devel
gb_est39:AW062071	+ 278.50	540.18	2.4e-21	570	! AW062071 660009C11.y1 660 - Mix
gb_gss13:AO916553	+ 278.00	540.18	4.5e-21	767	! AO916553 nbe0065020f CUGI Rice
gb_est48:AW720517	+ 277.00	542.03	3.5e-21	560	! AW720517 LjNEST18e3r Lotus japo
gb_est45:AW498397	+ 266.50	519.79	6.1e-20	598	! AW498397 660047D08.xl 660 - Mix
gb_gss10:AO687325	+ 257.50	500.53	7.2e-19	643	! AO687325 nxb0074M20r CUGI Rice
gb_est46:AW5666326	+ 257.00	513.55	1.4e-19	589	! AW5666326 660069F09.y1 660 - Mix
gb_est67:BE131390	+ 250.50	487.33	3.9e-18	507	! BE131390 L48-1365T3 Ice plant i
gb_est67:BE131520	+ 247.50	482.94	6.9e-18	512	! BE131520 L48-1504T3 Ice plant i
gb_est27:AI938086	+ 246.50	481.47	8.3e-18	489	! AI938086 SC41H08.x1 Gm-cl014 G
gb_est43:AW368451	+ 246.50	477.89	1.3e-17	653	! AW368451 CM4-HT0193-061099-022
gb_est29:AW094795	+ 246.00	475.81	1.7e-17	751	! AW094795 AU094795 Rice panicle
gb_est48:AW705316	+ 237.50	464.11	7.7e-17	451	! AW705316 SK59408.y1 Gm-cl019 G
gb_est45:AW508702	+ 235.00	457.78	1.7e-16	498	! AW508702 SI35406.y1 Gm-r1030 G
gb_est41:AW191191	+ 229.50	448.28	5.9e-16	433	! AW191191 687016F08.xl 687 - Ear

gb_est74:BE658710 - 224.50 437.88 2.2e-15 440 ! BE658710 GM700007A10F10 Gm-r
 gb_est39:AW055606 + 223.50 440.92 1.5e-15 292 ! AW055606 660009C11.xl 660 -
 gb_est74:BE658133 - 222.50 426.26 9.9e-15 809 ! BE658133 GM700005A10A5 Gm-r1
 gb_est73:BE640059 + 219.00 423.63 1.4e-14 562 ! BE640059 946046B05.y1 946 -
 seq_name: gb_est71:BE420587

seq_documentation_block:
 LOCUS BE420587 1212 bp mRNA EST 24-JUL-2000
 DEFINITION HWM000.D12 ITEC HWM Barley Leaf Library Hordeum vulgare cDNA clone
 ACCSSION BE420587
 VERSION BE420587.1 GI:9418430
 KEYWORDS EST.

SOURCE
 ORGANISM Hordeum vulgare
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.
 1 (bases 1 to 1212)

REFERENCE
 AUTHORS Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier
 S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,
 Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P.,
 Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y.,
 Pecchioni,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M.,
 Sorrells,M., Warburton,M. and Wenzel,G.
 International Triticeae EST Cooperative (ITEC): Production of
 Expressed Sequence Tags for Species of the Triticeae
 Unpublished (2000)
 Contact: Herrmann RG
 Botanisches Institut der LMU
 Menzinger Str. 67, D-80638 Munchen GERMANY
 Fax: 49 30 171683

FEATURES
 source
 1..1212
 /organism="Hordeum vulgare"
 /cultivar="Barke"
 /db_xref="taxon:4513"
 /clone="HWM000.D12"
 /clone_lib="ITEC HWM Barley Leaf Library"
 /tissue_type="leaf"
 /dev_stage="14 day old"
 /note="Vector: pBluescriptSK(-); 850 bp average insert
 size."

BASE COUNT 296 a 294 c 276 g 346 t
 ORIGIN

alignment_scores:
 Quality: 1034.50 Length: 299
 Ratio: 3.748 Gaps: 3
 Percent Similarity: 92.308 Percent Identity: 72.241

alignment_block:
 US-09-271-584A-2 x BE420587 ..
 Align seg 1/1 to: BE420587 from: 1 to: 1212

100 LysLysGlnPhePheArgAspPheValThrIleMetLeuPheGlyAlaVal 116
 57 AGAGAACAAATTCCTCCGCAATTCATGACTATATACATTGTTGGCGGTAGT 106
 116 lGlyThrIleLeuSerCysThrIleIleSerLeuGlyValThrGlnPheP 133
 107 TGGGACCTTGATCTCTCCAGTATATATACCTTGGTCCATGGGGCTAG 156
 133 heLysLeuAspIleGlyThrPheAspLeuGlyAspTyrIleuAlaIle 149
 157 TATCAAGCTGAACATAGCGCCCTTGAGCTTGAGACTACCTCGCACTT 206
 150 GlyAlaIlePheAlaAlaThrAspSerValCysThrLeuGlnValLeu 166

```
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
207 GGGCAATATTCGGCAAGCGACTCTCTCGCACCTTGCAGGTGTAAAG 256
166 nInAspGluThrProLeuLeuTyrSerLeuValPheGlyGluGlyVal 183
:|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
257 CCAAGATGAGACACTTCTTGACAGTTGGTGGTGGTGAAGGTGTG 306
183 alAsnAspAlaThrSerValValPheAsnAlaIleGlnSerPheAsp 199
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
307 TTAACGATGGACATCAGTTGTTGTTCAATGCAATCCAGAACTTGAT 356
200 LeuThrHisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLe 216
||| :|||||.....:|||||.....:|||||.....:|||||.....:
357 CTTCGAATTCAGTAGCCCAATTCCTACAATTCATTGGAATTCCT 406
216 uTyrLeuPheLeuLeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIleS 233
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
407 CTATCTATTTGGGCCCGCAGTACCTTCTGGAGTAGTCTAGTGGACTCTCA 456
233 erAlaTyrValIleLysLysLeuTyrPheGlyArgHisSerThrAspArg 249
457 GTGCTTATGTCATCAAGAACTGTACTTTGGCAGGCACTCCACTGATCGT 506
250 GluValAlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaG1 266
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
507 GAAGTTCGTATTATGATGCTCATGGCTTATTATCTTACATGCTGGCTGA 556
266 uLeuPheAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValM 283
557 ATTGCTTTGATTGAGTGGTATTCACGGTTTCTTCTGGTATTGTAA 606
283 etSerHisTyrThrTrpHisAsnValThrGluSerSerArgIleThrThr 299
607 TGTGCGACTATACCTGGCACAATGTACACAGAGTTCACGGGTCAACACC 656
300 LysHisThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLe 316
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
657 AAGCATGCCCTTGGCCACATTCATTCATCTCGAGACGTTCTCTTTCT 706
316 uTyrValGlyMetAspAlaLeuAspIleAspLysTyrArgSerValSerA 333
707 CTATGTTGGCATGGATGCAATGGATATAGAGAAAGTGGAAAATTGGTAGT 755
333 spThr.....ProGlyThrSerIleAlaValSerSerIleLeuMetGly 347
||| ||| |||.....:|||||.....:|||||.....:|||||.....:
756 A. ACATATAGCCCAATCAATCTATTGCTTGCCTGAGCTCCATTATTTGGCG 805
348 LeuValMetValGlyArgAlaAlaPheValPheProLeuSerPheLeuSe 364
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
806 TTGCTGCTGGTTGCAAGAGCTGCATTTGTTTCCCTCTATCT...ATCTC 852
364 r. AsnLeuAlaLysLysAsnGlnSerGluLysIleAsnPheAsnMetGln 380
| :|||||.....:|||||.....:|||||.....:|||||.....:
853 TCCAAATGACCAAAAACCTCCAGCGAGAGATCTCTGTAGGCGACAA 902
381 ValValIleTrpTrpSerGlyLeuMetArgGlyAlaValSerMet 395
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
903 GTTATTATTTGGTGGCG. GGGCTCATGAGAGGGCGCGTGTCAATG 946
```

seq_name: gb_est80:C99909

```
seq_documentation_block: 651 bp mRNA EST 21-SEP-1999
LOCUS C99909
DEFINITION C99909 Arabidopsis thaliana library (Motohashi R) Arabidopsis
thaliana cDNA clone 65, mRNA sequence.
ACCESSION C99909
VERSION C99909.1 GI:4714029
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsis.
```

```
REFERENCE 1 (bases 1 to 651)
AUTHORS Motohashi,R., Shinozaki,K. and Yamaguchi-Shinozaki,K.
TITLE Arabidopsis thaliana YAC C1C3B1 region specific cDNA,clone 2
JOURNAL Unpublished (1999)
COMMENT Contact: Motohashi R
Laboratory of Plant Molecular Biology
Institute of Physical and Chemical Research(RIKEN)
3-1-1, Koyadai, Tsukuba 305-0074, Japan
Tel: +81-298-36-4359
Email: motohashi@tc.riken.go.jp.
FEATURES
source
1..651
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="65"
BASE COUNT 161 a 156 c 152 g 178 t 4 others
ORIGIN
```

```
alignment_scores:
Quality: 1028.00 Length: 220
Ratio: 4.849 Gaps: 4
Percent Similarity: 96.364 Percent Identity: 95.909
alignment_block:
US-09-271-584A-2 x C99909 ..
```

Align seg 1/1 to: C99909 from: 1 to: 651

```
314 IlePheLeuTyrValGlyMetAspAlaLeuAspIleAspLysTrpArgSe 330
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
1 ATTCTTCTGATGTTGGAATGATGCTTGGACATTGACAAGTGGAGATC 50
330 rValSerAspThrProGlyThrSerIleAlaValSerSerIleLeuMetG 347
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
51 CGTGTGAGTCACACACCGGGAACATCGATCGAGTGCCTCAATCTTAATGG 100
347 lLeuValMetValGlyArgAlaAlaPheValPheProLeuSerPheLeu 363
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
101 GTCGTGTCATGTTGGAGAGAGCGGTTGCTCTTTCCGTTATCGTTTCTA 150
364 SerAsnLeuAlaLysLysAsnGlnSerGluLysIleAsnPheAsnMetG1 380
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
151 TCTAACTTAGCCCAAGAAGAAATCAAGCGAGAAAATCAACTTTAATGCA 200
380 nValValIleTrpTrpSerGlyLeuMetArgGlyAlaValSerMetAlaL 397
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
201 GGTGTGATTTGGTGGTCTGGTCTCATGAGAGTGTCTGTATCTATGGCTC 250
397 euAlaTyrAsnLysPheThrArgAlaGlyHisThrAspValArgGlyAsn 413
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
251 TTGCATACACAAGTTTACAGGGCGGGGCACACAGATGTACGGCGGAAT 300
414 AlaIleMetIleThrSerThrIleThrValCysLeuPheSerThrValVa 430
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
301 GCAATCATGATCAGGATACGATAACTGTCTCTCTTTTAGCACAGTGGT 350
430 lPheGlyMetLeuThrLysProLeuIleSerTyrLeuLeuProHisGlnA 447
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
351 GTTGTGATGTGTCGACCAACCACTCATAGCTACCTATTACCGCCACAGA 400
447 snAlaThrThrSerMetLeuSerAspAsnThrProLysSerIleHis 463
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
401 AGCCACCAACAGCATGTTATCTGATGACACACACCCCAAAATCCATACAT 450
464 lIleProLeuLeuAspGlnAspSerPheIleGluProSerGlyAsnHisAs 480
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
451 ATCCCTTTTGTGGACCAAGACCG.TTCATTGAGCGCTTCAGGGAACACAA 499
480 nValProArgProAspSerIleArgGlyPheLeuThrArgProThrArgT 497
|||||.....:|||||.....:|||||.....:|||||.....:|||||.....:
500 TGTGCTCGGCTGACAGTATACGGTCTTCTTGACACGGSCCCTCSAA 549
```

497 hrValHisTyrTrp.ArgGlnPheAspSerPheMetArgProVa 513
 |||||
 550 CCGTGGATTACTACTGGGAGACAAATTTGATGACTCTCTCGGACCGGT 599
 |||||
 513 lPhe.GlyGlyArgGlyPheValProPheVal.Pro.GlySerProThrG 529
 |||||
 600 CTTGGGAGGTGGCTTTGACCTTTGTTTCCARGTTTCCTCAACTG 649
 529 lu 529
 ||
 650 AR 651

seq_name: gb_est47:AW685820

seq_documentation_block: 653 bp mRNA EST 15-JUN-2000
 LOCUS AW685820 NF035804NR1F1000 Nodulated root Medicago truncatula cDNA clone
 DEFINITION NF035804NR 5', mRNA sequence.

ACCESSION AW685820

VERSION AW685820.1 GI:7560556

KEYWORDS EST.

SOURCE barrel medic.

ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
 Fabales; Fabaceae; Papilionoideae; Medicago.

REFERENCE 1 (bases 1 to 653)

AUTHORS Watson,B.S., Shin,H.-S., Lopez-Meyer,M., Scott,A.D., Harris,A.R.,
 Gonzales,R.A., Bell,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May
 ,G.D. and Paiva,N.L.

TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation

JOURNAL Medicago truncatula nodulated root library

COMMENT Unpublished (2000)

Contact: Paiva NL

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7317

Fax: 580 221 7380

Email: nlpaiva@noble.org

Insert Length: 653 Std Error: 0.00

Plate: 035 row: E column: 04

Seq primer: TCACACAGGAACAGCTATGAC.

Location/Qualifiers

FEATURES

source

1..653

/organism="Medicago truncatula"

/db_xref="taxon:3880"

/clone="NF035804NR"

/tissue_type="root"

/dev_stage="Pooled developmental"

/note="Vector: Lambda Zap; Four-week old Rhizobium

meliloti-inoculated Medicago truncatula roots, containing

a mixture of young and old roots and nodules."

BASE COUNT 167 a 133 c 148 g 205 t

ORIGIN

alignment_scores:

Quality: 714.50 Length: 218

Ratio: 3.841 Gaps: 2

Percent Similarity: 85.321 Percent Identity: 64.220

alignment_block:

US-09-271-584A-2 x AW685820 ..

Align seg 1/1 to: AW685820 from: 1 to: 653

307 SerPheLeuAlaGluThrPheLeuTyrValGlyMetAspAlaLe 323

|||||

2 TCATTGCTGCTGAGATATTTATCTTCTTATGTTGGGATGGATGCACT 51

323 uAspIleAspLysTrpArgSerValSerAspThrProGlyThrSerIleA 340
 |||||
 52 ACATATGGAGAAGTGGCGATTCGTAAGTCAAGGCCAAAAAATCAATTG 101
 |||||
 340 laValSerSerIleLeuMetGlyLeuValMetValGlyArgAlaAlaPhe 356
 ::|||
 102 GGGTCAGTTCATTGCTGTGGCAGCTATCTAGTGGGAAGAGCTGCATTT 151
 |||||
 357 ValPheProLeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGl 373
 |||||
 152 GTTTCCCTCTTCTTCATATCCAACTGTCTAAGAAGTCTCAATCTGA 201
 |||||
 373 ulysIleAsnPheAsnMetGlnValValIleTrpSerGlyLeuMeta 390
 |||||
 202 GAAATTTGAGTTTAAAGCAACAGTGAACAATTTGGTGGGCTGCTCATGC 251
 |||||
 390 rGClValAlaValSerMetAlaLeuAlaTyrAsnLysPheThrArgAlaGly 406
 ::|||
 252 GTGGAGCTGTTCTATGTCAGCTGTCTTACAACTGTTTACAGACTAGGA 301
 |||||
 407 HisThrAspValArgGlyAsnAlaIleMetIleThrSerThrIleThrVa 423
 |||||
 302 CATACTAAATTCGCTGAGAAATGCCATCATGATCACTAGTACTATTACAGT 351
 |||||
 423 lCysLeuPheSerThrValValPheGlyMetLeuThrLysProLeuIles 440
 |||||
 352 TGTACTCTTTTAGCACAGTGGTGTGGGTGATGACAAAGCCACTTGTGA 401
 |||||
 440 eTyrLeuLeuPro...HisGlnAsnAlaThrThrSerMetLeuSerAsp 455
 |||||
 402 GGTATTGTCACCTCTTCTTAAACACATAATAATCAGCATACCGTCCCG 451
 |||||
 456 AspAsnThrProLysSerIleHisIleProLeuLeuAspGlnAspSerPh 472
 |||||
 452 CCATCTAGCCCGAAATCAATTCAGTGTGCGACTTCTTGCAATGGAGAA.. 499
 |||||
 472 eIleGluProSerGlyAsnHisAsnValProArgProAspSerIleArgG 489
 |||||
 500GATGTAGGGGCAATGGTGGTACCCAAAGACCAAGAGATGGCGTA 545
 |||||
 489 lypheLeuThrArgProThrArgThrValHisTyrTrpArgGlnPhe 505
 ::|||
 546 CGCTCCTTAGGATTCCTAGTCATGGGTGATCATCACTACTGGCGCAAAATTT 595
 |||||
 506 AspAspSerPheMetArgProValPheGlyGlyArgGlyPheValProPh 522
 |||||
 596 GATGATTCCTTCATGGACCTGCTTTGGTGGCGGGGCTTTGTACCTTA 645
 |||||
 522 eVal 523
 :|||
 646 TGT 649

seq_name: gb_est71:BE440835

seq_documentation_block:

LOCUS BE440835

DEFINITION sp42f03.y1 Gm-cl043 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:

Gm-cl043-918 5', similar to TR:004655 004655 SIMILAR TO

SODIUM/HYDROGEN EXCHANGER. ; mRNA sequence.

ACCESSION BE440835

VERSION BE440835.1 GI:9440328

KEYWORDS EST.

SOURCE soybean.

ORGANISM Glycine max

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;

Fabales; Fabaceae; Papilionoideae; Glycine.

REFERENCE 1 (bases 1 to 502)

AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corvett,V., Khanna

,A., Holla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,

Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk

,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann

LINE	SEQUENCE	FEATURES
251	TCCAAAGCTTTTGACCTCAACCAAAATCGACCCCTTCAATTCCTTGGGCACTTT	300
212	LeuGlyAsnPhleuTyrLeuPheLeuLeuSerThrLeuLeuGlyAlaAla	228
301	TTGGGAAATTTCTTGAATCTATTTATTCGACCAATGCTTGAGAGTGT	350
228	aThrGlyLeuIleSerAlaTyrValIleLysLysLeuTyrPheGlyArgH	245
351	GACAGGTCTACTTAGTGTCTACATTATTAAGCTGTACATTCGACGC	400
245	isSerThrAspArgGluValAlaLeuMetMetLeuMetAlaTyrLeuSer	261
401	ACTCTACAGATCGGAGGTGCTTATGATTTAATGGCATACCTCTCC	450
262	TyrMetLeuAlaGluLeuPheAspLeuSerGlyIleLeuThrValPhePh	278
451	TACATGCTTCTGGAATTTGCTATCTGAGTGGCATCTCAGCTGATCTT	500
278	e 278	
501	T 501	
seq_name: gb_est27:AI938253		
seq_documentation_block:		
LOCUS	AI938253	EST
DEFINITION	sc41h02.y1 Gm-cl0104 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl0104-1708 5' similar to TR:004655 004655 SIMILAR TO SODIUM/HYDROGEN EXCHANGER. ;, mRNA sequence.	02-AUG-1999
ACCESSION	AI938253	
VERSION	AI938253.1	GI:5677123
KEYWORDS	EST.	
SOURCE	soybean.	
ORGANISM	Glycine max	
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.	
AUTHORS	Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.	
TITLE	Public Soybean EST Project	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com Possible reversed clone: similarity on wrong strand Seq primer: -40RP from Gibco High quality sequence stop: 381.	
FEATURES		
source	Location/Qualifiers	
1..502	/organism="Glycine max"	
	/db_xref="taxon:3847"	
	/clone="GENOME SYSTEMS CLONE ID: Gm-cl043-918"	
	/clone_lib="Gm-cl043"	
	/tissue_type="Hypocotyl and Plumule, germinating seeds"	
	/lab_host="DH10B"	
	/note="Vector: pT73Pac (Pharmacia); Site_1: EcoRI; Site_2: NotI; This cDNA library was constructed from mRNA isolated from hypocotyl and plumule tissues of seeds germinated for three days of the cultivar Williams. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a NotI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by digestion with EcoRI and NotI. The cDNA fragments were directionally cloned into the EcoRI-NotI restriction site of the pT73-Pac vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker."	
119 a	95 c 112 g 176 t	
BASE COUNT		
ORIGIN		
alignment_scores:		
Quality:	694.00	Length: 167
Ratio:	4.365	Gaps: 0
Percent Similarity:	95.210	Percent Identity: 81.437
alignment_block:		
US-09-271-584A-2 x BE440835 ..		
Align seg 1/1 to: BE440835 from: 1 to: 502		
112	LeuPheGlyAlaValGlyThrIleIleSerCysThrIleSerLeuGlu	128
1	TTGTTTGGTGTATTCGACATTAAATATCTTGACCATCAATCTTTGGG	50
128	yValThrGlnPhePheLysLysLeuAspIleGlyThrPheAspLeuGlyA	145
51	TGCCACACAAATTTTAAAGAGTTGGATGTTGTCACATGGAATTAGGGG	100
145	spTyrLeuAlaIleGlyAlaIlePheAlaIleThrAspSerValCysThr	161
101	ATTTCTAGCAATTTGGTGCATATTTCTGCAACGGATCTGTTGCACA	150
162	LeuGlnValLeuAsnGlnAspGluThrProLeuLeuTyrSerLeuValPh	178
151	TTGCAGGGGCTAAATCAGATGACACACCTTTGCTGTACAGCTTGTATT	200
178	eGlyGlyValValAlaAsnAspAlaThrSerValValValPheAsnAlaI	195
201	TGGGGAGGGTGTGTGAATGATGCTACATCAGTGGTCTTTTCAATGCAA	250
195	leGlnSerPheAspLeuThrHisLeuAsnHisGluAlaAlaPheHisLeu	211

plants. Complementary DNA was synthesized from mRNA using a 3' anchored poly (dT) primer. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by digestion with EcoRI and HindIII. The cDNA fragments were directionally cloned into the EcoRI-HindIII restriction site of the p77T3-Pac vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelidg.

BASE COUNT 122 a 101 c 99 g 164 t
ORIGIN

alignment_scores:

Quality: 684.00 Length: 160
Ratio: 4.442 Gaps: 0
Percent Similarity: 96.250 Percent Identity: 79.375

alignment_block:

US-09-271-584A-2 x A1938253 ..

Align seg 1/1 to: A1938253 from: 1 to: 486

```

274 LeuThrValPhePheCysGlyIleValMetSerHisTyrThrTrpHisAs 290
|||||
7 CTCACGTGATCTTTTGTGGATTGTCATGCTCATTTATACCTGCATAA 56

290 nValThrGluSerSerArgIleThrThrLysHisThrPheAlaThrLeuS 307
|||||
57 TGTGACTGAGAGTTCAAGATCACCACCAAGCATTTCTTTTGCAACCTGT 106

307 erPheLeuAlaGluThrPheIlePheLeuTyrValGlyMetAspAlaLeu 323
|||||
107 CTTTGTCTCAGATCTTTATCTTCCTTTATGTTGGTATGATGCCCTTG 156

324 AspIleAspLysTrpArgSerValSerAspThrProGlyThrSerIleAl 340
|||||
157 GACATTGAAAAATGGAATTTGTCAGTGATAGCCCTGGAACATCTGTAGC 206

340 aValSerSerIleLeuMetGlyLeuValMetValGlyArgAlaAlaPheV 357
|||||
207 ACTAGTTCAGTATTGTTGGGCTTAATTTCTTTGGAAGAGCAGCTTTTG 256

357 alPheProLeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGlu 373
|||||
257 TTTTCCCTTATCCTTCCTATCCACTTGGCTAAAAATCACCAAATGAG 306

374 LysIleAsnPheAsnMetGlnValValIleTrpTrpSerGlyLeuMetAr 390
|||||
307 AAAATCAGCTTCAGACAGCAAGATTATCATTTGTTGGCTGCTTATGAG 356

390 gGlyAlaValSerMetAlaLeuAlaTyrAsnLysPheThrArgAlaGlyH 407
|||||
357 AGGTCTCTTCAATCGGCACCTTGCATATAATACAGTTCCACCATGTCTG 406

407 IsThrAspValArgGlyAsnAlaIleMetIleThrSerThrIleThrVal 423
|||||
407 ACACCTTCACTAAGAAGCAATCAATCATGATGATCAACAGCACCATCAGT 456

424 CysLeuPheSerThrValValPheGlyMet 433
|||||
457 GTGCTTTTCAGCACAGCAGTGGTGTGTGCTG 486

```

seq_name: gb_est73:BE643915

seq_documentation_block:

LOCUS BE643915 469 bp mRNA EST
DEFINITION NCX1_048_H12_F NXCI (Nsf Xylem Compression wood Inclined) Pinus
taeda cDNA clone NXCI_048_H12 5', mRNA sequence.
ACCESSION BE643915
VERSION BE643915.1 GI:9956522
KEYWORDS EST.
SOURCE loblolly pine.

ORGANISM

Pinus taeda

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

REFERENCE

1 (bases 1 to 469)

AUTHORS

Sederoff, R.

TITLE

Molecular Basis of Wood Formation in the Pine Megagenome

JOURNAL

Unpublished (2000)

COMMENT

Contact: Johnson, Arthur

North Carolina State University

Tel: 919 515 7800

Fax: 919 515 7801

Email: ajohnson@unity.ncsu.edu

Seq primer: T3.

FEATURES

source

Location/Qualifiers
1..469
/organism="Pinus taeda"
/strain="Coastal plain loblolly pine from North Carolina"
/db_xref="taxon:352"
/clone="NXCI_048_H12"
/tissue_lib="NXCI (Nsf Xylem Compression wood Inclined)"
/tissue_type="Xylem"
/cell_type="Compression"
/dev_stage="Juvenile"
/lab_host="X11-Blue"
/note="Vector: BlueScript SK; Site_1: Eco RI; Site_2: XhoI
; The library is from early (spring) wood, taken from
three six-year old trees (three different genotypes), in
the juvenile phase. These trees were induced to form
compression wood by bending to a 45 degree angle and tying
them to the ground. Differentiating xylem was harvested
from the bottoms of the inclined stems, and a mixture of
all three genotypes was used for the library. oligo-dT
primed cDNA was directionally cloned into the EcoRI-XhoI
BlueScript SK vector arms. NOTE: The sequences contain a
'cDNA adapter' between the EcoRI site and the start of the
EST. The adapter sequence is 'AATTCGGCACGAG'."
BASE COUNT 109 a 91 c 97 g 155 t 17 others
ORIGIN

alignment_scores:

Quality: 632.00 Length: 156
Ratio: 4.514 Gaps: 0
Percent Similarity: 89.744 Percent Identity: 77.564

alignment_block:

US-09-271-584A-2 x BE643915 ..

Align seg 1/1 to: BE643915 from: 1 to: 469

```

213 GlyAsnPheLeuTyrLeuPheLeuLeuSerThrLeuLeuGlyAlaAla 229
|||||
2 GGTAACCTTTGTTTCATCTCTTTCTGACAAGCACTGCATTGGGGATTGCAGC 51

229 rGlyLeuIleSerAlaTyrValIleLysLysLeuTyrPheGlyArgHis 246
|||||
52 TGGCCTTTTAAAGTGATACATCATCAAAAGCTTTGTACTTTTGGGAAGGCAT 101

246 erThrAspArgGluValAlaLeuMetMetLeuMetAlaTyrLeuSerTyr 262
|||||
102 CTACTGATCGTGAAGTAGCAATTTATGATACTCATGCGGTATCTCTATAC 151

263 MetLeuAlaGluLeuPheAspLeuSerGlyIleLeuThrValPhePheCy 279
|||||
152 ATGCTCGCTGAGCTGTTTCACCTCAGTGGGATCTTGACTGTTTCTCTCTG 201

279 sGlyIleValMetSerHisTyrThrTrpHisAsnValThrGluSerSera 296
|||||
202 TGCATTGTTATGCTCTACCTACACTTGGCATATATGACTGAAAGCTCAA 251

296 rgIleThrThrLysHisThrPheAlaThrLeuSerPheLeuAlaGluThr 312
|||||
252 GGATTACCACCAAGCATGTCATTGCAACAATGTCATTATTCGCTGAACA 301

```

cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

313 PheIlePheLeuTyrValGlyMetAspAlaLeuAspIleAspIleTyrPar 329
|||||
302 TTTATCTTCCTTTTATGTTGGGAATGGATGCTCTGGACATTTGAAAATGGAA 351

329 gSerValSerAspThrProGlyThrSerIleAlaValSerSerIleLeuM 346
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
352 AATGGTGAAGAAGAGCCCTGGACATCTNTNNNTCAGCAGTGTCTCC 401

346 etGlyLeuValMetValGlyArgAlaPheValPheProLeuSerPhe 362
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
402 TGGGNTTGATCTTNTTGGANNNCNCNTTGTGTNNCACTTTCTTTC 451

363 LeuSerAsnLeuAlaLys 368
|||||
452 TTGTCAAATTCCTTAAAG 469

seq_name: gb_est76:BE803390

seq_documentation_block: 574 bp mRNA EST 20-SEP-2000
LOCUS BE803390
DEFINITION sr55g05.yl Gm-c1051 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-c1051-2001 5' similar to TR:Q9SXJ8 Q9SXJ8 OSNXXI PROTEIN. ;
mRNA sequence.

ACCESSION BE803390
VERSION BE803390.1 GI:10234502
KEYWORDS EST.

SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
Fabales; Fabaceae; Papilionoideae; Glycine.

REFERENCE 1 (bases 1 to 574)
AUTHORS Shoemaker, R., Kelm, P., Vodkin, L., Erpelding, J., Corvett, V., Khanna
, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
R., Waterston, R. and Willson, R.

TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com
Trace considered overall poor quality
High quality sequence stop: 1.

FEATURES
source Location/Qualifiers
1..574
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1051-2001"
/clone_lib="Gm-c1051"
/tissue_type="floral meristematic mRNA"
/lab_host="DH10B"

note=Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from floral
meristematic mRNA provided by Dr. Halina Knap of Clemson
University. Complementary DNA was synthesized from mRNA
using a primer consisting of a poly(dT) sequence with a
XhoI restriction site. EcoRI adapters were ligated to the
blunt-ended cDNA fragments followed by XhoI digestion. The
cDNA fragments were directionally cloned into the
EcoRI-XhoI restriction site of the pBluescript vector. The
ligated cDNA fragments were transformed into DH10B host

alignment_scores:
Quality: 590.50 Length: 192
Ratio: 3.601 Gaps: 4
Percent Similarity: 85.417 Percent Identity: 60.417

alignment_block:
US-09-271-584A-2 x BE803390 ..

Align seg 1/1 to: BE803390 from: 1 to: 574

345 LeuMetGlyLeuValMetValGlyArgAlaAlaPheValPheProLeuSe 361
|||||
2 TTGTTGGGTCTAAATCTTTTGGAGAGCAGCTTTGGTTTCCCTTATC 51

361 rPheLeuSerAsnLeuAlaLysLysAsnGlnSerGluLysIleAsnPheA 378
|||||
52 CTCCTCTCCAACTTGGCTAAATAATCACCAGAAATGAGAAATCAGCTTCA 101

378 snMetGlnValIleThrTrpSerGlyLeuMetArgGlyAlaValSer 394
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
102 GACAGCATGTTATCATTTTGGTGGGTGCTTATGAGAGGTGCTGTTCA 151

395 MetAlaLeuAlaTyrAsnLysPheThrArgAlaGlyHisThrAspVala 411
|||||
152 ATCGGCACTTGCATATAAATCAGTTCACCATGCTGGGCACACTTCACTAAG 201

411 gGlyAsnAlaIleMetIleThrSerThrIleThrValCysLeuPheSerT 428
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
202 AGCAATGCGATCATGATCACAAGCACCACATCATCTGTGCTTTTCAGCA 251

428 hrValValPheGlyMetLeuThrLysProLeuIleSerTyrLeuLeuPro 444
|||||
252 CAGTGGTGTAGTCTGTGTGACTAAGCCCACTTAAGGCTTTTACTGGCC 301

445HisGlnAsnAlaThrSerMetLeuSerAspAspAs 457
|||||
302 CATACTCCACATCATATAAGAATCAAGCATCAGATCACCACAGATCCATC 351

457 nThrPro.....LysSerIleHisIleProLeuLeuAsp.....GlnA 470
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
352 TACTCCAAGTACAAAAATCAGTCTCCACTTCTTGAGAGTGCCCAAG 401

470 spSerPheIleGluProSerGlyAsnHisAsnValProArgProAspSer 486
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
402 AATCTCAAGTCGATATGATGGC...CATGATATTCATCGTCCAAGCAGC 448

487 IleArgGlyPheLeuThrArgProThrArgThrValHisTyrTyrTrpAr 503
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
449 ATCCGTGCTTTACTTCCACTCCAACACACACTGTTTCATCGTGTGGCG 498

503 gGlnPheAspSerPheMetArgProValPheGlyGlyArgGlyPheV 520
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
499 TGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 548

520 alProPheValProGlySerProThr 528
|||||
549 TTCTATAGAACCTAGCTCACCAC 574

seq_name: gb_est36:AV408114

seq_documentation_block: 424 bp mRNA EST 23-MAY-2000
LOCUS AV408114
DEFINITION japonicus cDNA clone MWL036dl1_r 5', mRNA sequence.
ACCESSION AV408114
VERSION AV408114.1 GI:7720968
KEYWORDS EST.

SOURCE Lotus japonicus.
ORGANISM Lotus japonicus
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids 1; Fabales; Fabaceae; Papilionoideae; Lotus.

REFERENCE 1 (bases 1 to 424)
AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE Generation of 7137 non-redundant expressed sequence tags from a legume, Lotus japonicus
JOURNAL DNA Res. 7 (2), 127-130 (2000)
MEDLINE 20277479
COMMENT Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1..424
/organism="Lotus japonicus"
/db_xref="taxon:34305"
/clone="MML036d11r"
/dev_stage="young plants (two-week old)"
/note="vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI; isolate=Miyakojima MG-20"
XhoI: 103 a 78 c 95 g 148 t

BASE COUNT 103 a 78 c 95 g 148 t

ORIGIN
120 ILSerCysThrIleSerLeuGlyValThrGlnPheLysLysLe 136
1 ATATCTTGCACCATATACTTGGTGTCTAGCGAAGCTTTAAAGTGGAT 50
136 uAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleP 153
51 AGATGTGGTCCACTGGAGTGGGATTTACTAGCAATGGTGAATAT 100
153 heAlaIleThrAspSerValCysThrLeuGlnValLeuAsnGlnAspGlu 169
101 TTGCTGCTACAGATTCTGTTGCACATTCAGGTGTTAAATCAAGATGAG 150
170 ThrProLeuLeuTyrSerLeuValPheGlyGluGlyValValAsnAspAl 186
151 ACACCTTGTCTACAGTCTGTATTTGGGGAGGTGTGTCAATCATGTC 200
186 aThrSerValValPheAsnAlaIleGlnSerPheAspLeuThrHisL 203
201 TACATCAGTGGTCTTTCAATGCAATCCAAAGCTTTGACCTCAATCAA 250
203 euAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTyrLeuPhe 219
251 TTAATCTTCAATGCTTTGCACTTTTGGCAACTTCTGTATCTGTGTT 300
220 LeuLeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIleSerAlaTyrVa 236
301 ATCGCAAGCATTGCTGGGGTTTGGACAGGTCTTCTCAGTGGCTTACAT 350
236 lIleLysLysLeuTyrPheGlyArgHisSerThrAspArgGluValAlaL 253
351 CATTAAGAGCTATATGGCAGGCACTCTACAGATCGTGGTGGTCTC 400
253 euMetMetLeuMetAlaTyrLeu 260
|||||

401 TTATGATGCTAATGTCATACCTT 423
seq_name: gb_est71:BE433982

seq_documentation_block: 519 bp mRNA EST 24-JUL-2000
LOCUS BE433982
DEFINITION EST7405060 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA clone cLEG9M8, mRNA sequence.

ACCESSION BE433982
VERSION BE433982.1 GI:9431825
KEYWORDS EST.
SOURCE tomato.

ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids 1; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 519)
AUTHORS Alcalá, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E., Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Ronning, C.M., Niernman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.
TITLE Generation of ESTs from tomato fruit tissue, breaker stage
JOURNAL Unpublished (2000)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU

FEATURES
source
1..519
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEG9M8"
/clone_lib="tomato breaker fruit, TIGR"
/tissue_type="pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/note="vector: pBluescriptSKmCuadapt; Site_1: EcoRI; Site_2: XhoI; Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of the fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."
BASE COUNT 147 a 127 c 111 g 133 t 1 others

ORIGIN
alignment_scores:
Quality: 554.50 Length: 176
Ratio: 3.798 Gaps: 3
Percent Similarity: 82.955 Percent Identity: 62.500

alignment_block:
US-09-271-584A-2 x BE433982 ..
Align seg 1/1 to: BE433982 from: 1 to: 519

359 ProLeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGluLysI 375
1 CCATTATCACTATATCCAACTTAATGAAGAAATCTCGGAAGAAAAAT 50
375 eAsnPheAsnMetGlnValIleTyrPheSerGlyLeuMetArgGlyA 392
51 TACCTTTAGGCACCAAGTGATATATGTTGGCAGGTTTGTATGAGAGGG 100
392 laValSerMetAlaLeuAlaTyrAsnLysPheThrArgAlaGlyHisThr 408
101 CAGTGTCCATGGCACTGGCATATAAAGTTCACTCATGGGGGACACACT 150
409 AspValArgGlyAsnAlaIleMetIleThrSerThrIleThrValCysLe 425

```

151 CAACCTGCAGACATGCAATGATGATTACAGACACATACCACTTTCT 200
      ::::: |||||||
425 upheSerThrValPheGlyMetLeuThrTyrProLeuLeuSerTyrL 442
      |||||||:|||||:|||||:|||||:|||||:|||||:|||||
201 ATTCAGCACAAATGGTATTGCTCTAATGACAAAACCCCTTATAAGTCTCC 250
      ::::: |||||||
442 euLeuProHisGlnAsnAlaThrThrSerMetLeuSerAspAsnThr 458
      ||||||| ||::: ::::: |||
251 TGCTGCCACACAGAGCAATGAGTACAGTGCATCAGGTGCAATATCT 300
      |||||||:|||||:|||||:|||||:|||||:|||||
459 ProLysSerThrHisIleProLeuLeuAsp.....GlnAspSerPheII 473
      |||||||:|||||:|||||:|||||:|||||:|||||
301 CCTAAGTCTCTACAGCCCACTCTAGGAGTCGAGAGACTCT..... 345
      ::::: |||||||
473 eGluProSerGlyAsnHisAsnVal.....ProArgProAspSerI 487
      ::::: |||||||
346 .....GAAGGTGATTAAATGTCACAGATCTTCCTCAACCACCAAGTT 388
      ::::: |||||||
487 leArgGlyPheLeuThrArgProThrArgThrValHisTyrTyrTrpArg 503
      ::::: |||||||
389 TGAGGATGCTACTTACCGCACCAAGTCAATAAGTGCATCGTACTGGCGC 438
      ::::: |||||||
504 GlnPheAspSerPheMetArgProValPheGlyArgGlyPheVa 520
      ::::: |||||||
439 AAGTTTGACGATGATTCATCGCGCCCTATGTTGGTGGTGGGATTTGC 488
      ::::: |||||||
520 lProPheValProGlySerProThrGlu 529
      ::::: |||||||
489 TCCTNCTGCTCCTCGTCTCCACCGAA 516

```

seq_name: gb_est73:BE607262

```

seq_documentation_block: 401 bp mRNA EST 22-AUG-2000
LOCUS BE607262
DEFINITION NCXCI_034_G06_F NCXI (Nsf Xylem Compression wood Inclined) Pinus
      taeda cDNA clone NCXCI_034_G06 5', mRNA sequence.
ACCESSION BE607262
VERSION BE607262.1 GI:9884342
KEYWORDS EST.
SOURCE loblolly pine.
ORGANISM Pinus taeda
REFERENCE Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
      Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
AUTHORS 1 (bases 1 to 401)
TITLE Sederoff,R.
JOURNAL Molecular Basis of Wood Formation in the Pine Megagenome
COMMENT Unpublished (2000)
      Contact: Johnson, Arthur
      North Carolina State University
      Tel: 919 515 7800
      Fax: 919 515 7801
      Email: ajohnson@unity.ncsu.edu
Seq primer: T3.

```

```

FEATURES
      source
      1..401
      /organism="Pinus taeda"
      /strain="Coastal plain loblolly pine from North Carolina"
      /db_xref="taxon:3352"
      /clone="NCXCI_034_G06"
      /clone_lib="NCXI (Nsf Xylem Compression wood Inclined)"
      /tissue_type="Xylem"
      /cell_type="Compression"
      /dev_stage="Juvenile"
      /lab_host="Xll-Blue"
      /note="Vector: Bluescript SK; Site_1: Eco RI; Site_2: XhoI
      ; The library is from early (spring) wood, taken from
      three six-year old trees (three different genotypes), in
      the juvenile phase. These trees were induced to form
      compression wood by bending to a 45 degree angle and tying
      them to the ground. Differentiating xylem was harvested
      from the bottoms of the inclined stems, and a mixture of
      all three genotypes was used for the library. oligo-dT

```

primed cDNA was directionally cloned into the EcoRI-XhoI
 Bluescript SK vector arms. NOTE: The sequences contain a
 'cDNA adapter' between the EcoRI site and the start of the
 EST. The adapter sequence is 'AATTCGGCAGAG'."

BASE COUNT 98 a 84 c 79 g 125 t 15 others
 ORIGIN

alignment_scores: Quality: 544.00 Length: 133
 Ratio: 4.496 Gaps: 0
 Percent Similarity: 90.977 Percent Identity: 77.444

alignment_block:
 US-09-271-584A-2 x BE607262 ..

Align seg 1/1 to: BE607262 from: 1 to: 401

```

213 GlyAsnPheLeuTyrLeuPheLeuSerThrLeuLeuGlyAlaAlaLath 229
      |||||||:|||||:|||||:|||||:|||||:|||||
2 GGTAACTTTGTTTCATCTCTTCTGACAAGCACTGCATTTGGGATTGCAGC 51
      ::::: |||||||
229 rGlyLeuIleSerAlaTyrValIleLysLysLeuTyrPheGlyArgHiss 246
      :|||||:|||||:|||||:|||||:|||||:|||||
52 TGGGCTTTTAAGTGCATACATCATCAAGCTTTGTACTTTGGAAGGCAC 101
      ::::: |||||||
246 erThrAspArgGluValAlaLeuMetMetLeuMetAlaTyrLeuSerTyr 262
      |||||||:|||||:|||||:|||||:|||||:|||||
102 CTACTGATGCTGAAGTAGCAATATGATACTCATGGCGTATCTCTCATAC 151
      ::::: |||||||
263 MetLeuAlaGluLeuPheAspLeuSerGlyIleLeuThrValPhePheCy 279
      |||||||:|||||:|||||:|||||:|||||:|||||
152 ATGCTCGCTGAGCTGTTTCACCTGAGTGGGATCTTGACTNNTTTCTCTN 201
      ::::: |||||||
279 sGlyIleValMetSerHisTyrThrTroHisAsnValThrGluSerSera 296
      |||||||:|||||:|||||:|||||:|||||:|||||
202 NGGCATTNNNATGTCCTCACTACACTGGCATATGCTGACTGAAGCTCNN 251
      ::::: |||||||
296 rgIleThrThrLysHisThrPheAlaThrLeuSerPheLeuAlaGluThr 312
      |||||||:|||||:|||||:|||||:|||||:|||||
252 NNATTACCAACCAAGCATGCAATTTGCAACAATGTCATTTATCGCTGAAACA 301
      ::::: |||||||
313 PheIlePheLeuTyrValGlyMetAspAlaLeuAspIleAspLysTrpAr 329
      |||||||:|||||:|||||:|||||:|||||:|||||
302 TTTATCTTCTTTATGTGGGAATGGATGCTCTGGACATTTGAAAAATGGAA 351
      ::::: |||||||
329 gSerValSerAspThrProGlyThrSerIleAlaValSerSerIleLeu 345
      :|||||:|||||:|||||:|||||:|||||:|||||
352 ATGGTGAAAGAGAGCCCTNNCACATCTGTTNNCCTCAGCAGTGTCTC 400

```

seq_name: gb_est80:C91832

```

seq_documentation_block: 378 bp mRNA EST 20-APR-1998
LOCUS C91832
DEFINITION E31686_1A, mRNA sequence.
      Rice panicle shorter than 3cm Oryza sativa cDNA clone

```

```

ACCESSION C91832
VERSION C91832.1 GI:3061200
KEYWORDS EST.
SOURCE Oryza sativa.
ORGANISM Oryza sativa

```

```

REFERENCE Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
      Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
AUTHORS 1 (bases 1 to 378)
TITLE Rice cDNA from panicle
JOURNAL Sasaki,T. and Yamamoto,K.
      Unpublished (1997)
COMMENT Contact: Takuji Sasaki
      National Institute of Agrobiological Resources
      Rice Genome Research Program
      2-1-2 Kannondai,Tsukuba
      Ibaraki,
      Japan 305

```


Tel: 0298-38-7441

Fax: 0298-38-7468

Email: tsasaki@affrc.go.jp.

FEATURES

Location/Qualifiers

source

1..378

/organism="Oryza sativa"

/strain="Nipponbare"

/db_xref="taxon:4530"

/clone="E31686_1A"

/dev_stage="rice panicle shorter than 3cm"

/note="Organ: panicle"

BASE COUNT 88 a 74 c 89 g 121 t 6 others

ORIGIN

alignment_scores:

Quality: 522.00 Length: 125

Ratio: 4.424 Gaps: 2

Percent Similarity: 94.400 Percent Identity: 83.200

alignment_block:

US-09-271-584A-2 x C91832

Align seg 1/1 to: C91832 from: 1 to: 378

238 LysLysLeuTyrPheGlyArgHisSerThrAspArgGluValAlaLeuMe 254

2 AAGAGCTATACATTGGGAGCGATTCTACTGACCGTGAGGTGCCCTAT 51

254 tMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPheAspLeu 271

52 GATGCTCATGCTTACCTTTCATATATGCTGGCTGAGTTGCTAGATTGA 101

271 erGlyLeuThrValPhePheCysGlyIleValMetSerHisTyrThr 287

102 GCGGATTCTCACCCTATCTCTGTGGTATGTTGAATGTACATTACACT 151

288 TrpHisAsnValThrGluSerSerArgIleThrThrLysHisThrPheAl 304

152 TGGCATAACGTCACAGAGAGTTCAAGAGTTACACAAGACCGATTGGC 201

304 aThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrValGlyMet 321

202 AACTCTGTCTTCATTGCTGAGACTTTTCTCTCTCTCTCTCTCTCTCT 251

321 spAlaLeuAspIleAspLysTrpArgSerValSerAspThrProGlyThr 337

252 ATGCTTGGATTGAAAATGGGAGNTTCCAGTGACAGACCTGGNAAA 301

338 SerIleAlaValSerSerIleLeuMet.GlyLeuVal.MetValGlyArg 353

302 TCCATTNGGGTAAGCTCAATTGCTAGGATGGTTCCCTGATTGGAGN 351

354 AlaAlaPheValPheProLeu 360

352 GCTGCTTTTGGNAATTCCCTCG 372

seq_name: gb_est48:AW694621

seq_documentation_block:

LOCUS AW694621 757 bp mRNA EST 15-JUN-2000

DEFINITION NF078C11ST1F1085 Developing stem Medicago truncatula cDNA clone

ACCESSION AW694621

VERSION AW694621.1 GI:7569383

KEYWORDS EST.

SOURCE barrel medic.

ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.

REFERENCE 1 (bases 1 to 757)

AUTHORS

He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon
R.A.

TITLE

Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula stem library

JOURNAL

Unpublished (2000)

COMMENT

Contact: Dixon RA

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7302

Fax: 580 221 7380

Email: radixon@noble.org

Insert Length: 757 Std Error: 0.00

Plate: 078 row: C column: 11

Seq primer: TCACACAGGAAACAGCTATGAC.

FEATURES

source

1..757

/organism="Medicago truncatula"

/db_xref="taxon:3880"

/clone="NF078C11ST"

/clone_lib="Developing stem"

/tissue_type="stem"

/dev_stage="Pooled developmental"

/note="Vector: Lambda Zap; Contains a mixture of

internodal stem segments"

BASE COUNT 230 a 171 c 143 g 213 t

ORIGIN

alignment_scores:

Quality: 496.50 Length: 167

Ratio: 3.598 Gaps: 4

Percent Similarity: 82.635 Percent Identity: 61.677

alignment_block:

US-09-271-584A-2 x AW694621

Align seg 1/1 to: AW694621 from: 1 to: 757

334 ThrProGlyThrSerIleAlaValSerSerIleLeuMetGlyLeuValMe 350

2 AGTCCTGGAACATCTATAGCAGCAAGTTCAGTATGTTGGGCTAATACT 51

350 tValGlyArgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeu 367

52 TCTTGAAGAGCAGCGCTTTGTTTCCCTTATCCTTCTTATCCAACTTGA 101

367 laLysLysAsnGlnSerGluLysIleAsnPheAsnMetGlnValValIle 383

102 CTAAAAAATCAACAATCAGACATCTCCTTCAGACAGCAAGTGATCAAT 151

384 TrpTrpSerGlyLeuMetArgGlyAlaValSerMetAlaLeuAlaTyrAs 400

152 TGGTGGCTGGTCTTATGAGAGTGCTGTTTCAATGGCCTTGGCTATAA 201

400 nLysPheThrArgAlaGlyHisThrAspValArgGlyAsnAlaIleMetI 417

202 TCAGTTCCACCATGTCGGGGCATACTCAACTGCTAGCAATGCAATCATGA 251

417 leThrSerThrIleThrValCysLeuPheSerThrValValPheGlyMet 433

252 TAACACAGCACCATCACTGTTGTCCTTTTCAGCACAGTGGTGGTTGG 301

434 LeuThrLysProLeuIleSerTyrLeuLeuProHisGlnAsnAlaThrTh 450

302 CTGACTAAGCCACTTATAAGGCTTCTACTACCTCATCTCTTAAATCACAAG 351

450 rSerMet....LeuSerAspAsnThrProLysSerIleHis..... 463

352 CAGCATGACACACACAGATCTACTACTCCAAAATACATTTCATTCCCA 401

464IleProLeuLeuAspGlnAsp.....SerPheIle 473

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 542 1805

Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.

Seq primer: JEN REV

High quality sequence stop: 516

POLYA=No.

Location/Qualifiers

FEATURES

source

1..556

/organism="Sorghum bicolor"

/db_xref="taxon:4558"

/note="Organ: Anthracnose induced 1 (P11)"

/clone_lib="Pathogen induced 1 (P11)"

Two-week-old sorghum plants 48 hr after inoculation;

Vector: pBluescript II from Lambda Zap II; Site.1: XhoI;

Site.2: EcoRI; Two-week-old sorghum plants (PTX 623

cultivar) were infected with pathogen (isolate FRM421 of

Colletotrichum graminicola, which is a sorghum isolate).

RNA was prepared from infected leaves harvested from 45

seedlings 48 hours after inoculation. Note: young

seedlings (2 weeks old) exhibit juvenile resistant

reaction, which is an incompatible interaction. As they

grow older (4 weeks or older), plants resume susceptibility

to anthracnose disease. The library was made from poly-A

RNA in the cloning vector lambda ZAP II. Clones to be

sequenced were prepared by mass excision. WARNING: While

most or all ESTs are expected to derive from the host

plant, no effort was made to eliminate ESTs deriving from

the pathogen." 104 a 158 c 147 g 147 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 463.00 Length: 121

Ratio: 4.097 Gaps: 0

Percent Similarity: 93.388 Percent Identity: 72.727

alignment_block:

US-09-271-584A-2 x BE600656 ..

Align seg 1/1 to: BE600656 from: 1 to: 556

8 LysLeuProSerLeuSerThrSerAspHisAlaSerValAlaLeuAs 24

||||| :|||:|||||:|||||: :|||:|||||:|||||:|||||:|||||

188 AAATCCGGCGCTCTCGGTCGCGACTAGGACGCATCGTCCATCAA 237

24 nLeuPheValAlaLeuLeuCysAlaCysIleValLeuGlyHisLeuG 41

||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||

238 CATCTTCATCGCGCTCTCGGTCGCGACTAGGACGCATCGTCCATCAA 287

41 luGluAsnArgTrpMetAsnGluSerIleThrAlaLeuIleGlyLeu 57

||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||

288 AGGGAACGATGGTGAAGACGAGTCCATCACCAGCGCTGTGATGGTCTC 337

58 GlyThrGlyValThrIleLeuLeuLeuSerLysGlyLysSerSerHisLe 74

||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||

338 ATCACCAGGAGCGTGATCTGCTGCTGCTCACTAATGGACAACTCACGCAT 387

74 uLeuValPheSerGluAspLeuPhePheIleTrpLeuLeuProIleI 91

||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||

388 TCTTGTGTTCAGCGAGGACCTGTTTTCATATACCTGCTTCGCGCATAA 437

91 lePheAsnAlaGlyPheGlnValLysLysGlnPhePheArgAsnPhe 107

||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||

438 TCTTCAATCCGCGTTCAGGTAAGTAAGAAAGCAATCTTCGCCAATTT 487

108 ValThrIleMetLeuPheGlyAlaValGlyThrIleIleSerCysThrI 124

||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||

488 ATAACAATTATTGTTTGGTGTGTTGGAACCTCTGATTCTCTTCTTAAT 537

124 eIleSerLeuGly 128

|||||:|||||:|||||:|||||:|||||:|||||:|||||

538 AATCTCTCTGGT 550

seq_name: gb_est27:AI941126

seq_documentation_block:

LOCUS AI941126 352 bp mRNA EST 13-DEC-1999

DEFINITION SB84h02.y1 Gm-cl010 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:

Gm-cl010-1708 5' similar to TR:004635 004635 SIMILAR TO

SODIUM/HYDROGEN EXCHANGER. ; mRNA sequence.

ACCESSION AI941126

VERSION. AI941126.1 GI:5688111

KEYWORDS EST.

SOURCE soybean.

ORGANISM Glycine max

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;

Fabales; Fabaceae; Papilionoideae; Glycine.

REFERENCE 1 (bases 1 to 352)

Shoemaker,R., Kelm,P., Vodkin,L., Erpelnding,J., Coryell,V., Khanna

,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,

Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk

,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann

,R., Waterston,R. and Wilson,R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: Genome Systems, Inc. 4633 World

Parkway Circle St. Louis, Missouri 63134 For further information

call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)

427-3324 or contact: clones@genomesystems.com or

info@genomesystems.com web site: www.genomesystems.com

Trace considered overall poor quality

Possible reversed clone: similarity on wrong strand

High quality sequence stop: 1.

FEATURES

source

1..352

/organism="Glycine max"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl010-1708"

/clone_lib="Gm-cl010"

/tissue_type="young cotyledons of greenhouse grown plants"

/dev_stage="2cm long 12 week old"

/lab_host="Xl10-Gold"

/note="Vector: pBluescript II SK+; Site.1: EcoRI; Site.2:

XhoI; This cDNA library was constructed from mRNA isolated

from immature cotyledons (100-200mgs) of old greenhouse

grown plants. The cDNA library was prepared using the

Stratagene pBluescript II SK(+) library construction kit.

Complementary DNA was synthesized from mRNA using a primer

consisting of a poly (dT) sequence with a XhoI restriction

site. EcoRI adapters were ligated to the blunt-ended cDNA

fragments followed by XhoI digestion. The cDNA fragments

were directionally cloned into the EcoRI-XhoI restriction

site of the pBluescript vector. The ligated cDNA fragments

were transformed into Xl10-Gold host cells. This library

was constructed by Dr. Lila Vodkin and Dr. Anu Khanna."

BASE COUNT 82 a 65 c 85 g 120 t

ORIGIN

alignment_scores:

Quality: 442.00 Length: 113

Ratio: 4.131 Gaps: 0

Percent Similarity: 94.690 Percent Identity: 76.106

alignment_block:

US-09-271-584A-2 x AI941126

Align seg 1/1 to: AI941126 from: 1 to: 352

```
114 GlyAlaValGlyThrIleIleSerCysThrIleIleSerLeuGlyValTh 130
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
5  GGTGCTATTGCCACATTAATATCTGGTACCATCATACTTTGGGTGCCAC 54

130 rGlnPhePheLysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrL 147
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
55  ACAAAATTTTAAGAGGTTGGATGTTGGTCCACTGGAATTAGGGGATTTC 104

147 euAlaIleGlyAlaIlePheAlaIleThrAspSerValCysThrLeuGln 163
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
105 TAGCATTTGGTGCAATATTGCTGCACGGATTCTGTTGCACATTGCAG 154

164 ValLeuAsnGlnAspGluThrProLeuLeuTyrSerLeuValPheGlyG1 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
155 GTGCTTAATCAGGATGACACACCTTTGCTGTACAGTCTTGTTTGGGA 204

180 uGlyValValAsnAspAlaThrSerValValPheAsnAlaIleGlnS 197
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
205 GGGTGTGTGAATGATGCTACATCAGTGGTCTTTTCAATGCAATCCAGA 254

197 erPheAspLeuThrHisLeuAsnHisGluAlaAlaPheHisLeuLeuGly 213
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
255 GCTTGACCTCAACCAATCGACCCCTCAATCGCTGGGCACCTTTTGGGA 304

214 AsnPheLeuTyrLeuPheLeuLeuSerThrLeuLeuGly 226
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
305 AATTCTGTATCTATTATTATGCAAGCACAAATGCTTGGGA 343
```


Align seg 1/1 to: Z22591 from: 1 to: 2178

```
1 MetLeuAspSerLeuValSerLysLeuProSerLeuSerThrSerAspHi 17
286 ATGTTGGATTCTCTAGTGTGCGAACCTGCCCTCGTTATCGACATCTGATCA 335
17 sAlaSerValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34
336 CGCTTCGTGGTTCGGTTGAATCTCTTTGTTGGACATCTCTTTGTGCTTGTGA 385
34 leValLeuGlyHisLeuLeuGluGluAsnArgTrpMetAsnGluSerIle 50
386 TTGTTCTTGGTCACTCTTTTGAAGAGAAATAGATGATGAACGAATCCATC 435
51 ThrAlaLeuLeuLeuGlyLeuGlyThrGlyValThrIleLeuLeuIleSe 67
436 ACCCCCTGTGATTTGGGCTAGGCACCTGGTTTACCATTGTTGTGATTAG 485
67 rLysGlyLysSerSerHisLeuLeuValPheSerGluAspLeuPhePheI 84
486 TAAAGGAAAAGCTCGCATCTTCCTCGCTTTAGTGAAGATCTTTTCTTCA 535
84 leftyLeuLeuProProIleIlePheAsnAlaGlyPheGlnValLysLys 100
536 TATATCTTTTCCCAACCATATATTCGAATGACAGGTTTCAAGTAAAAAAG 585
101 LysGlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaValG1 117
586 AAGCAGTTTTTCCCAATTCGTCGACATATATGCTTTTGGTGTGCTGTGG 635
117 yThrIleIleSerCysThrIleLeuSerLeuGlyValThrGlnPhePheL 134
636 GACTATTATTTCTTGGCAACATCATATCTTAGGTGTAACACAGTCTTTTA 685
134 ysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleGly 150
686 AGAAGTTGGACATTTGGAAACCTTTGACTTGGGTGATTATCTTGTCTATTGGT 735
151 AlaIlePheAlaIaThrAspSerValCysThrLeuGlnValLeuAsnG1 167
736 GCCATATTTGTCACACAGATTACATCTAGGTGTAACAGTCTTGAATCA 785
167 nAspGluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValVala 184
786 AGACGAGACACCTTTGCTTTACAGCTTTGTATTTCGAGAGGGTGTGTGA 835
184 snAspAlaThrSerValValPheAsnAlaIleGlnSerPheAspLeu 200
836 ATGATGCAACGTCAGTTGTGCTTTCACGCGGATTCAGAGCTTTGTATCTC 885
201 ThrHisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTy 217
886 ACTCACCTTAACACGAGAGCTGCTTTCACTCTTGGAACTCTTGTGA 935
217 rLeuPheLeuLeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIleSerA 234
936 TTTGTTTCTCCTAAAGTACCTTGTGCTTGGTGTGCAACCGGCTGTGATAAGTG 985
234 laTyrValIleIleLysLysLeuTyrPheGlyArgHisSerThrAspArgGlu 250
986 CGTATGTTATCAAGAAGCTATACTTTTGAAGGCACCTCACTGACCGGAG 1035
251 ValAlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLe 267
1036 GTTGCCCTTATGATGCTTATGCGGTACTTCTTATATGCTTGTGAGCT 1085
267 upPheAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMetS 284
1086 TTTTCGACTTGACGGGTATCCCTCACTGTGTTTTTCTGTGGTATTGTGATGT 1135
284 erHisTyrThrTrpHisAsnValThrGluSerSerArgIleThrThriys 300
|||||
```

```
1136 CCCATTACACATGGCACAAATGTAAACGGAGAGCTCAAGAAATAACACAAAG 1185
301 HisThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTy 317
|||||
1186 CATACCTTTTGAACCTTTGTCATTTCTTCGGGAGACATTTATTTCTTGTA 1235
317 rValGlyMetAspAlaLeuAspIleAspLysTrpArgSerValSerAspT 334
|||||
1236 TGTGTGAATGGATGCTTGGCATTGGACAAGTGGAGATCCGTGAGTGACA 1285
334 hrProGlyThrSerIleAlaValSerSerIleLeuMetGlyLeuValMet 350
1286 CACCGGGAACATCGATCGAGTGCAGTGCATCTCAATCGGCTGGTGCATG 1335
351 ValGlyArgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAl 367
|||||
1336 GTTGAAGAGCAGCGTTCGTCTTTCCGTTATCGTTTCTATCTAACTTAGC 1385
367 aLysLysAsnGlnSerGluLysIleAsnPheAsnMetGlnValValIleT 384
|||||
1386 CAAGAAGAATCAAGCGAGAAATCAACTTTAAACATGCAGGTGTGTGATTT 1435
384 rPTrpSerGlyLeuMetArgGlyAlaValSerMetAlaLeuAlaTyrAsn 400
1436 GGTGTGCTGTCTCATGAGAGGTGCTGTATCTATGCGCTCTTGCATACAC 1485
401 LysPheThrArgAlaGlyHisThrAspValArgGlyAsnAlaIleMetI1 417
|||||
1486 AAGTTTACAAGGGCGGCGCACACAGATGTACGCGGAATGCAATCATGAT 1535
417 ehThrSerThrIleThrValCysLeuPheSerThrValValPheGlyMetL 434
1536 CACGAGTAGCATAACTGTCTGCTTTTTAGCACAGTGGTGTGTGTATGC 1585
434 euThrLysProLeuIleSerTyrLeuLeuProHisGlnAsnAlaThrThr 450
1586 TGACCAACCACTCATAGCTACCTATTACCGCACAGAACGCCACACAG 1635
451 SerMetLeuSerAspAspAsnThrProLysSerIleHisIleProLeuLe 467
|||||
1636 AGCATGTTTATCTGATGACAAACACCCCAAAATCCATACATATCCCTTTGT 1685
467 uAspGlnAspSerPheIleGluProSerGlyAsnHisAsnValProArgP 484
1686 GGACCAAGACTCGTTTCATTGAGCCTTCAGGGAACCAACATGTGCCTCGC 1735
484 roAspSerIleArgGlyPheLeuThrArgProThrArgThrValHisTyr 500
1736 CTGACAGTATACGTGCTTCTTGACACGGCCCTCGAACCGTGCATTAC 1785
501 TyrTrpArgGlnPheAspAspSerPheMetArgProValPheGlyGlyAr 517
1786 TACTGGAGACAAATTTGATGACTCTCTTCATGCGACCGCTCTTTTGGAGTGC 1835
517 gGlyPheValProPheValProGlySerProThrGluArgAsnProProA 534
1836 TGCTTTGTACCCCTTTGTTCCAGGTTCTCCAATGAGAGAAACCCCTCCNG 1885
534 spLeuSerLysAla 538
1886 ATCTTAGTAAGGCT 1899
seq_name: /SID56/gcgdata/geneseq/geneseq/NA1999.DAT:22595
seq_documentation_block:
ID Z22595 standard; cDNA; 2284 BP.
XX
AC Z22595;
XX
DT 18-JAN-2000 (first entry)
XX
DE Arabidopsis thaliana Na/H transporter gene.
XX
```

KW Sodium; proton; antiport; transporter; salt tolerance; salt management;
 KW transgenic plant; survival; soil; farming; accumulation; irrigation;
 KW crop; ss.

XX Arabidopsis thaliana.

XX WO9947679-A2.

XX 23-SEP-1999.

XX 18-MAR-1999; 99WO-CA00219.

XX 18-MAR-1998; 98US-0078474.

XX 15-JAN-1999; 99US-0116111.

XX (BLUM/) BLUMWALD E.

XX (APSE/) APSE M.

XX (SNED/) SNEDDEN W.

XX (AHAR/) AHARON G.

XX Blumwald E, Apse M, Snedden W, Aharon G;

XX WPI; 1999-571840/48.

XX P-PSDB; Y40905.

XX Nucleic acid molecules encoding sodium/proton transport polypeptides,
 useful in genetic engineering salt tolerance in crop plants

XX Claim 48; Fig 5A-B; 93pp; English.

XX The invention relates to an isolated nucleic acid molecule encoding
 a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment
 and capable of increasing salt tolerance in a cell. This sequence
 corresponds to the gene encoding a transporter from Arabidopsis thaliana.
 CC The Na/H transporter polypeptides provide a means of intracellular salt
 management, particularly in plants. The sequences are useful for
 CC producing transgenic plants that are capable of surviving in soil with
 CC high salt levels that would normally inhibit growth of the crop species.
 CC This would be useful in farming land in areas that are generally
 CC considered unproductive through salt accumulation and poor irrigation,
 CC e.g. in India, Australia, and prairies in USA or Canada. Commercial
 CC crops, such as potatoes, tomatoes, brassica, cotton, sunflower,
 CC strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley,
 CC atriplex, sorghum, alfalfa, salicornia and others would benefit from
 CC increased salt tolerance.

XX SQ Sequence 2284 BP; 589 A; 444 C; 485 G; 766 T; 0 other;

alignment_scores:

Quality: 2720.00 Length: 540
 Ratio: 5.065 Gaps: 2
 Percent Similarity: 99.444 Percent Identity: 99.444

alignment_block:

US-09-271-584a-2 x 222595

Align seg 1/1 to: 222595 from: 1 to: 2284

1 MetLeuAspSerLeuValSerLysLeuProSerLeuSerThrSerAspHi 17

401 ATCTGGATTCTCTAGTGTGGAACCTGCTTGTATCGACATCTGATCA 450

17 sAlaSerValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34

451 CGCTTCTGTGGTGGCTTCTTGTGTCACCTTCTTGTGCTTGTGTA 500

34 leValLeuGlyHisLeuLeuGluGluAsnArgTrpMetAsnGluSerile 50

501 TTGTTCTTGGTCATCTTTTGGAGAGATAGATGGATGAACCAATCCATC 550

51 ThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeuLeuIleSe 67

551 ACCGCTTGTGATTTGGTGGCTAGGCACCTGGTCTTACCATTTTTGTGATTAG 600
 67 rLysGlyLysSerSerHisLeuLeuValPheSerGluAspLeuPhePheI 84
 601 TAAAGGAAAAAGCTCGCATCTTCTCGTCTTTTAGTGAAGATCTTTTCTTCA 650
 84 leTyrLeuLeuProProIleIlePheAsnAlaGlyPheGlnValLysLys 100
 651 TATATCTTTTGGCACCATTATATCAATGACGGTTTCAAGTAAAAAAG 700
 101 LysGlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaValGI 117
 701 AAGCAGTTTTTCCGCAATTTCTGCTGACTATTATGCTTTTGGTGTGTGG 750
 117 yThrIleIleSerCysThrIleIleSerLeuGlyValThrGlnPhePheI 134
 751 GACTATTATTCTTGCACAATCATATCTTAGGTGTAAACACAGTCTTTA 800
 134 yLysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleGly 150
 801 AGAAGTTGGACATTGGAACCTTTGACTTTGGGTGATTATCTTGTCTATTGGT 850
 151 AlaIlePheAlaAlaThrAspSerValCysThrLeuGlnValLeuAsnGI 167
 851 GCCATATTGTGTCACAGATTCAGTATGACATGCAGGTTCTGANTCA 900
 167 nAspGluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValVala 184
 901 AGACGAGACACCTTTGCTTTTACAGTCTTGTATTCGGAGAGGGTGTGTGA 950
 184 snAspAlaThrSerValValValPheAsnAlaIleGlnSerPheAspLeu 200
 951 ATGATGCAACGTCAGTGTGGTCTTCAACGGGATTCAGAGCTTTGATCTC 1000
 201 ThrHisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTy 217
 1001 ACTCACCTAAACACGAAAGCTGCTTTTTCATCTTCTTGGAAACTTCTGTGA 1050
 217 rLeuPheLeuLeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIleSerA 234
 1051 TTTGTTTCTCCTAAGTACCTTGTGGTGTGCAACCGCTGTGATAAGTG 1100
 234 laTyrValIleLysLysLeuTyrPheGlyArgHisSerThrAspArgGlu 250
 1101 CGTATGTTATCAAGAGCTATACTTTGGAAGGCATCAACTGACCGACAG 1150
 251 ValAlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLe 267
 1151 GTTGCCCTTATGATGCTTATGCGGTATCTTCTTATATGCTTGTGAGCT 1200
 267 uPheAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMets 284
 1201 TTTTCGACTTGAAGCGGTATCTCCTCCTGCTTTTCTGTGTGTTGTGATGT 1250
 284 erHisTyrThrTrpHisAsnValThrGluSerSerArgIleThrLys 300
 1251 CCATTACACATGGCAATGTAACGGAGCTCAAGATACACAAAG 1300
 301 HisThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTy 317
 1301 CATACCTTTGCAACTTTGTCAATTTCTTCTGCGAGACATTTATTTCTTCTGA 1350
 317 rValGlyMetAspAlaLeuAspIleAspLysTrpArgSerValSerAspT 334
 1351 TGTGGAATGGATGGCTTGGACATTTGACAAGTGGAGATCCCGTGGTGAACA 1400
 334 hrProGlyThrSerIleAlaValSerSerIleLeuMetGlyLeuValMet 350
 1401 CACCGGGAACATCGATCGCAGTGAAGTCAATCTTAATGGGTCTGTGTCATG 1450
 351 ValGlyArgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAl 367
 1451 GTTGGGAAGACGACGCTTCTTCTTCCGTTATGCTTCTATCTAATTAACCTAGC 1500

367	aLysLysAsnGlnSerGluLysIleAsnPheAsnMetGlnValValIleT	394
1501	CAGAAGATCAAGCGAGAAATCAATTTAAACATGCAGGTTGTGATTT	1550
384	rpTpSerGlyLeuMetArgGlyAlaValSerMetAlaLeuAlaTyAsn	400
1551	GGTGGTCTCATGAGAGTGCTGATCTATCTATGGCTTCATGCATACAAC	1600
401	LysPheThrArgAlaGlyHisThrAspValArgGlyAsnAlaIleMetII	417
1601	AACTTTACAAGGCGCGGCACACAGATCTACCGGGAATGCAATCATGAT	1650
417	eThrSerThrIleThrValCysLeuPheSerThrValValPheGlyMetL	434
1651	CACGAGTACGATAACTGTCTGCTCTTTTAGCACAGTGGTGTGGTATGC	1700
434	euthrLysProLeuIleSerTyrrLeuLeuProHisGlnAsnAlaThrThr	450
1701	TGACCAAAACCATCATTAAGCTACCTATTACCGCACCAAGCCACCAACG	1750
451	SerMetLeuSerAspAsnThrProLysSerIleHisIleProLeuLe	467
1751	AGCATGTTATCTGATGACAAACACCCAAAATCCATACATATCCCTTGT	1800
467	uAspGlnAspSerPheIleGluProSerGlyAsnHisAsnValProArgP	484
1801	GGACCAAGACTCGTTCATTGAGCCCTTCAGGGAACCAATGTGCCTGGC	1850
484	roAspSerIleArgGlyPheLeuThrArgProThrArg.ThrValHisTy	500
1851	CTGACAGTACGTGGCTTCTTTGACACGGCCCACTCGAACCGTGCATTA	1900
500	r.TyrTrpArgGlnPheAspAspSerPheMetArgProValPheGlyGly	516
1901	CTAACTGGAGACAATTTGATGACTCTTTCATGCGACCCGCTTTTGGAGT	1950
517	ArgGlyPheValProPheValProGlySerProThrGluArgAsnProPr	533
1951	CGTGGCTTTGTACCTCTTGTCTCCAGGTCTCCAACTGAGAGAAACCTCC	2000
533	oAspLeuSerLysAla 538	
2001	TGATCTATTAGAGCT 2016	
seq_name: /SID6/gcgdata/geneseq/geneseqn/NA1999.DAT:222592		
seq_documentation block:		
ID	222592 standard; CDNA; 1788 BP.	
XX		
AC	222592;	
XX		
DT	18-JAN-2000 (first entry)	
XX		
DE	Arabidopsis thaliana Na/H transporter gene AtNHX2.	
XX		
KW	Sodium; proton; antiport; transporter; salt tolerance; salt	
KW	transgenic plant; survival; soil; farming; accumulation; ir	
KW	crop; ss.	
OS	Arabidopsis thaliana.	
XX		
PN	W09947679-A2.	
XX		
PD	23-SEP-1999.	
XX		
PF	18-MAR-1999; 99WO-CA00219.	
XX		
PR	18-MAR-1998; 98US-0078474.	
XX		
PR	15-JAN-1999; 99US-0116111.	
XX		
PA	(BLUM/) BLUMWALD E.	
PA	(APSE/) APSE M.	

```
(S/NED/) SNEDDEN W.  
PA      PA  
XX      AHARAK) AHARON G.  
  
PI      Blumwald E, Apse M, Snedden W, Aharon G;  
XX  
XX      WPJ; 1999-571840/48.  
DR      P-PSDB; Y40902.  
XX  
XX      Nucleic acid molecules encoding sodium/proton transport polypeptide useful in genetic engineering salt tolerance in crop plants  
PT  
FT  
XX  
XX      Claim 4; Fig 1B; 93pp; English.

The invention relates to an isolated nucleic acid molecule capable of increasing salt tolerance in a plant Na/H antiport (PNHX) transporter polypeptide, or a fragment thereof, which encodes a protein having at least one amino acid substitution compared to the gene encoding the AtNHX2 transporter from Arabidopsis thaliana. The Na/H transporter polypeptides provide a means of intracellular pH regulation in plants. The sequences are used to produce transgenic plants that are capable of surviving under conditions of increased salinity by inhibiting growth of the host plant. This would be useful in farming land in areas that are generally considered unproductive through salt accumulation and poor soil quality, e.g., in India, Australia, and prairies in USA or Canada. CC crops, such as potatoes, tomatoes, brassica, cotton, sunflower, strawberries, spinach, lettuce, rice, soybean, corn, wheat, sorghum, alfalfa, salicornia and others would benefit from increased salt tolerance.



Sequence 1788 BP; 475 A; 367 C; 370 G; 576 T; 0 other;


```


CC species. This would be useful in farming land in areas that are generally
 CC considered unproductive through salt accumulation and poor irrigation,
 CC e.g. in India, Australia, and prairies in USA or Canada. Commercial
 CC crops, such as potatoes, tomatoes, brassica, cotton, sunflower,
 CC strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley,
 CC atriplex, sorghum, alfalfa, salicornia and others would benefit from
 CC increased salt tolerance.
 XX
 SQ Sequence 378 BP; 88 A; 74 C; 89 G; 121 T; 6 other;

alignment_scores:
 Quality: 522.00 Length: 125
 Ratio: 4.424 Gaps: 2
 Percent Similarity: 94.400 Percent Identity: 83.200

alignment_block:

US-09-271-584A-2 x 222606 ..

Align seg 1/1 to: 222606 from: 1 to: 378

238 LysLysLeuTyrPheGlyArgHisSerThrAspArgGluValAlaLeuMe 254
 |||||
 2 AAGAAGCTATACATTGGAAGGCATTCTACTGACCGTGAGGTTGCCCTTAT 51
 |||||
 254 tMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPheAspLeuS 271
 |||||
 52 GATGCTCATCGCTTACCTTCATATATGCTGGCTGAGTGTAGATTGA 101
 |||||
 271 erGlyLeuLeuThrValPhePheCysGlyLeValMetSerHisTyrThr 287
 |||||
 102 GCGCAATCTCACCGTATCTCTGTGGTATTGTAATGTCACATTACACT 151
 |||||
 288 TrpHisAsnValThrGluSerSerArgIleThrThrLysHisThrPheAl 304
 |||||
 152 TGGCATAACCTCACAGAGTTCAAGAGTTACAAAGACGACCATTTGC 201
 |||||
 304 aThrLeuSerPheLeuAlaGluThrPheLeuTyrValGlyMetA 321
 |||||
 202 AACTCTGCTCTCATGCTGAGACTTTCTCTCTGATGTGGATGG 251
 |||||
 321 spAlaLeuAspIleAspLysTrpArgSerValSerAspThrProGlyThr 337
 |||||
 252 ATGCATTGGATATTGAAAAATGGGAGNTTNCACAGACACCTGGNAA 301
 |||||
 338 SerIleAlaValSerSerIleLeuMet.GlyLeuValMetValGlyArg 353
 |||||
 302 TCCATTNGGTAAGCTCAATTTGCTAGGAGTTGGTTCCTGATTGGAAGN 351
 |||||
 354 AlaAlaPheValPheProLeu 360
 |||||
 352 GCTGCTTTTGNAAATCCCTTG 372

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:222602

seq_documentation_block:

ID 222602 standard; cDNA; 1683 BP.

XX

AC 222602;

XX

DT 18-JAN-2000 (first entry)

XX

DE Arabidopsis thaliana Na/H transporter gene AtNHX4 variant #1.

XX

KW Sodium; proton; antiport; transporter; salt tolerance; salt management;

KW transgenic plant; survival; soil; farming; accumulation; irrigation;

KW crop; ss.

XX

OS Arabidopsis thaliana.

XX

PN W09947679-A2.

XX

PD 23-SEP-1999.

XX 18-MAR-1999; 99WO-CA00219.
 PF
 PR 18-MAR-1998; 98US-0078474.
 PR 15-JAN-1999; 99US-0116111.
 XX
 PA (BLUM/) BLUMWALD E.
 PA (APSE/) APSE M.
 PA (SNED/) SNEDDEN W.
 PA (AHAR/) AHARON G.
 XX
 PI Blumwald E, Apse M, Snedden W, Aharon G;
 XX
 DR WPI: 1999-571840/48.
 DR P-PSDB; Y40906.
 XX
 PT Nucleic acid molecules encoding sodium/proton transport polypeptides,
 PT useful in genetic engineering salt tolerance in crop plants
 XX
 PS Claim 4; Fig 1D; 93pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule encoding a
 CC plant Na/H antiport (PNHX) transporter polypeptide, or a fragment and
 CC capable of increasing salt tolerance in a cell. This sequence corresponds
 CC to one of 2 splice variants of the AtNHX4 transporter gene from
 CC Arabidopsis thaliana. The Na/H transporter polypeptides provide a means
 CC of intracellular salt management, particularly in plants. The sequences
 CC are useful for producing transgenic plants that are capable of surviving
 CC in soil with high salt levels that would normally inhibit growth of the
 CC crop species. This would be useful in farming land in areas that are
 CC generally considered unproductive through salt accumulation and poor
 CC irrigation, e.g. in India, Australia, and prairies in USA or Canada.
 CC Commercial crops, such as potatoes, tomatoes, brassica, cotton,
 CC sunflower, strawberries, spinach, lettuce, rice, soybean, corn, wheat,
 CC rye, barley, atriplex, sorghum, alfalfa, salicornia and others would
 CC benefit from increased salt tolerance.
 XX
 SQ Sequence 1683 BP; 392 A; 360 C; 367 G; 564 T; 0 other;
 alignment_scores:
 Quality: 512.50 Length: 536
 Ratio: 1.567 Gaps: 21
 Percent Similarity: 61.007 Percent Identity: 30.410
 alignment_block:
 US-09-271-584A-2 x 222602 ..
 Align seg 1/1 to: 222602 from: 1 to: 1683
 18 AlaSerValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34
 |||||
 37 GCGGCGCGGTGGTATCTGCTTCAGATTATGATGCTCGTCTTCCTT 86
 |||||
 34 eValLeuGlyHisLeuLeuGluGluAsnArg.....TrpMetAsnGluS 49
 |||||
 87 CGTCTCGGCATGCTCCGCGGTATCGATTCCACTATCTCTCTGAAG 136
 |||||
 49 erIleThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeuLeu 65
 |||||
 137 CCAGCGGTTCGCTCTCATTTGTTTAAATCGTGGTATA...CTTGCTAAT 183
 |||||
 66 IleSerLysGlyLysSerSer.....HisLeuLeuValPheSerGluAs 80
 |||||
 184 ATCTCCGATAGTACACTAGCATTAGGACGTGGTTTAAATTTCCACGAGA 233
 |||||
 80 pLeuPheHeiLeuLeuLeuProIleIlePheAsnAlaGlyPheG 97
 |||||
 234 GTCTCTCTCTGTTTGTGCTCCCATCATATTCAGTCAGGTTTCA 283
 |||||
 97 InValLysLysLysGlnPhePheArgAsnPheValThrIleMetLeuPhe 113
 |||||
 284 GTCTTCAACCTAAACCATTCCTTTTAACTTTGGAGCCATTGTTACCTTT 333

```
114 GlyAlaValGlyThrIleIleSerCysThrIle.....IleSe 126
   : : : : : : : : : : : : : : : : : : : : : : : :
334 GCTATCATCGGAACCTTTTGGCTTCAGTTGTACTGGTGGTCTGGTTA 383
126 rLeuGlyValThrGlnPhePheLysLysLeuAspIleGlyThrPheAspL 143
   ||||| : : : : : : : : : : : : : : : : : : : : : :
384 TCTTGGCGGCTCTATGATCTC.....ATGTATAAATTCCTCT 421
143 euGlyAspTyrLeuAlaIleGlyAlaIlePheAlaAlaThrAspSerVal 159
   : : : : : : : : : : : : : : : : : : : : : : : :
422 TTGTTGAGTGTCTTATGTTTGGTGCACTTATATCAGCTACGGACCTGTC 471
160 CysThrLeuGlnValLeuAsnGlnAspGluThrProLeu...LeuTyrSe 175
   : : : : : : : : : : : : : : : : : : : : : : : :
472 ACTGTACTCTATATTCAGGATGGGCACCGATGTTAACTGTATGTC 521
175 rLeuValPheGlyGluGlyValValAsnAspAlaThrSerValValP 192
   ||||| : : : : : : : : : : : : : : : : : : : : : :
522 TTTGGTCTTTGGAGATCATGTTCTGAATGATGCTATGGCAATATCATTTG 571
192 heAsnAlaIleGlnSerPheAspLeuThrHisLeuAsnHisGluAla 208
   : : : : : : : : : : : : : : : : : : : : : : : :
572 ACAGAACAAATGCTTAGTAACCCGCTGCTCTGCTGGGGAACATTTT 621
209 PheHisLeuGlyAsnPheLeuTyrLeuPheLeuLeuSerThrLeuLe 225
   ||| : : : : : : : : : : : : : : : : : : : : : :
622 TTTCTGTTGGTGCATCAGGTTTTCAGACTTTTGTGCTCAATCTCTGCC 671
225 uGlyAlaAlaThrGlyLeuIleSerAlaTyrValIleLysLysLeuTyrP 242
   ||||| : : : : : : : : : : : : : : : : : : : : : :
672 AGGGGTTGGGTTGGATTTCACCTTCAGCTTACTCTTTAAG.....TATG 715
242 heGly.....ArgHisSerThrAspArgGluValAlaLeuMetMet 255
   ||| : : : : : : : : : : : : : : : : : : : : : :
716 CAGGATTTGGACACCGAGAACTTCAGAACTTGGAGTGTGCTCTCTTTGTA 765
256 LeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPheAspLeuSerG 272
   ||||| : : : : : : : : : : : : : : : : : : : : : :
766 CTTTCCCGTATTTTTCATACATGCTTCGACAGAGGTGTGGTCTCTCCGG 815
272 yIleLeuThrValPhePheCysGlyIleValMetSerHisTyrThrP 289
   ||||| : : : : : : : : : : : : : : : : : : : : : :
816 CATTTCTTCTATCTACTCTTCACAGGAATGTTATGAAGCGCTACACTTCT 865
289 isAsnValThrGluSerSerArgIleThrThrLysHisThrPheAlaThr 305
   ||| : : : : : : : : : : : : : : : : : : : : : :
866 CAATCTCTCAGAACTTCACAGAGTTTCGTATCTCTTTTTCACCTG 915
306 LeuSerPheLeuAlaGluThrPheIlePheLeuTyrValGlyMetAsp... 321
   ||||| : : : : : : : : : : : : : : : : : : : : : :
916 ATATCTTCGCTAGCAGAACTTTCAGTTCATTTACATGGGATTTGATAT 965
322 .AlaLeuAspIleAspLysTyrArgSerValSerAspThrProGlyThr 338
   ||||| : : : : : : : : : : : : : : : : : : : : : :
966 TCCCATGGAGCAGCATAGCTGGTCCCAT..... 993
338 erIleAlaValSerSerIleLeuMetGlyLeuVal..... 349
   ||||| : : : : : : : : : : : : : : : : : : : : : :
994 .....GTTGGGTTTATCTTTTCTATGTTATCTCTCTATTTACTGAT 1035
350 .....MetValGlyArgAlaAlaPheValPheProLeuSerPheLe 363
   : : : : : : : : : : : : : : : : : : : : : : : :
1036 GGTGATGATTTGTCAGTGGCTCTCAATGATTTTGGTGTGCTATTTT 1085
363 uSerAsnLeuAlaLysLysAsnGlnSerGluLysIleAsnPheAsnMetG 380
   ||||| : : : : : : : : : : : : : : : : : : : : : :
1086 GGTCAACCTATTATTAGACAG...GAGAACAGAGATACCTATGAAGCACC 1132
380 InValValIleThrTrpSerGlyLeuMetArgGlyAlaValSerMetAla 396
   || : : : : : : : : : : : : : : : : : : : : : : : :
1133 AAAAGGCCCTTTGGTATATGAGACTT...CGAGGGGCAATGGCAATTTGCA 1179
```

```
397 LeuAla.TyrAsnLysPheThrArgAlaGlyHisThrAspValArgGlyA 413
   ||||| : : : : : : : : : : : : : : : : : : : : : :
1180 CTTGCATCTCAATCACTTATGATGATCTACAGAGGGTCTACGGCCAAATCAT 1229
413 snAlaIleMetIleThrSerThrIleThrValCysLeuPheSerThrVal 429
   : : : : : : : : : : : : : : : : : : : : : : : :
1230 CT.....TTACTGCAAAACCACTATTTGTTGTCCACGGTTTTA 1270
430 ValPheGlyMetLeuThrLysProLeuIleSerTyrLeuProHisG 446
   : : : : : : : : : : : : : : : : : : : : : : : :
1271 CTAATAGGAGGTTGCAGAGGTAATAATCTTGAAGCTTTG..... 1309
446 nAsnAlaThrThrSerMetLeuSerAspAsnThrProLysSerIleH 463
   : : : : : : : : : : : : : : : : : : : : : : : :
1310 .....GAAGTTGTAGTGACGAT..... 1327
463 isIleProLeuLeuAspGlnAspSerPhe.....IleGluProSer 476
   ||||| : : : : : : : : : : : : : : : : : : : : : :
1328 .....CTTGAT...GACTCCATGCTGAAGGCTTTGAAGAGAGC 1363
477 GlyAsnHisAsnValProArgProAspSerIleArgGlyPheLeuThrAr 493
   : : : : : : : : : : : : : : : : : : : : : : : :
1364 GATCATGATGATGCTCCTCCTCTTTAGCATTTGGAGCTTCATCTGACGA 1413
493 gProThr.....ArgThrValH 499
   : : : : : : : : : : : : : : : : : : : : : : : :
1414 GGATACATCATCATCAGAGAGCAGGTTCAAGATGAAGCTGAAGGAGTTTC 1463
499 isTyrTyrTrpArgGlnPhe.....AspAspSerPheMetArgPro 512
   || : : : : : : : : : : : : : : : : : : : : : : : :
1464 ACAAAACCACTACATCATTCACCGCGTTGCACAAAAAACTTTCGACTCCG 1513
513 ValPhe 514
   |||
1514 TTCTTC 1519
seq_name: /SID56/gcgdata/geneseq/geneseq/NA1997.DAT.T51330
seq_documentation_block:
ID T51330 standard; cdna; 2584 BP.
XX
AC T51330;
XX
DT 29-APR-1997 (first entry)
XX
DE Human Na+/H+ exchanger isoform NHE3 composite cDNA.
XX
KW NHE3; Na+/H+ exchanger; sodium ion-proton exchanger; antiporter;
KW chronic metabolic acidosis; chronic renal failure; hypertension;
KW diabetic nephropathy; familial diarrhoea; cystic fibrosis; kidney;
KW small intestine; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 11..2515
FT /*tag= a
FT /transl_except= pos:1970..1972;aa:Arg
FT /note= "codon CAC (bases 1870-1972) codes for Arg"
XX
PN W09703196-A1.
XX
PD 30-JAN-1997.
XX
PF 10-JUL-1996; 96WO-US11670.
XX
PR 11-JUL-1995; 95US-0001061.
XX
PA (MAGA-) MAGAININ PHARM INC.
XX
PI Brant SR, Donowitz M, Tse C, Yun CHC;
XX
DR WPI; 1997-119051/11.
```

DR P-PSDB; W11325.
 XX New DNA encoding the human NHE3 ion exchange protein - used to
 PT screen for agents that affect NHE3 activity, e.g. for treatment of
 PT chronic metabolic acidosis or hypertension
 XX
 XX Claim 1; Fig 3; 47pp; English.
 XX
 CC A composite cDNA sequence (T51330) codes for human NHE3 (W11325), a
 CC characteristic Na⁺/H⁺ ion exchanger of the brush border of the
 CC kidney and small intestine. It is composed of the previously known
 CC partial clones HKC-3 and HKC-5, and of newly isolated clone 23-3.
 CC The latter was isolated from human kidney RNA by RT-PCR (see also
 CC T51331-32). Full-length cDNA can be prep'd. using human NHE3
 CC specific primers (see also T51333-34). Transformed host cells
 CC expressing NHE3 can be used to identify agents that affect human
 CC NHE3. Such agents may be useful for treating diseases of the
 CC kidney and gastro-intestinal tract.
 XX
 SQ Sequence 2584 BP; 455 A; 844 C; 791 G; 494 T; 0 other;

alignment_scores:

Quality: 463.00 Length: 510
 Ratio: 1.470 Gaps: 18
 Percent Similarity: 61.765 Percent Identity: 28.627

alignment_block:

US-09-271-584A-2 x T51330 ..

Align seg 1/1 to: T51330 from: 1 to: 2584

20 ValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysIleValLe 36
 179 GTCAATCGCGCTCGATCCTCGTGGCCAGCTTGCCCAAG.....AT 219
 36 uGlyHisLeuLeuGluGluAsnArgTrpMetAsnGluSerIleThrAlaL 53
 220 CGGTTCCACCTGTCCACAGGTCACCGGTGGTTCGCCGAGAGCGCC 269
 53 euLeuIleGlyLeuGlyThrGlyValThrIleLeuLeuIleSerLysGly 69
 270 TGCTCATCGTCTGGGC.....CTGGTCTGGGCGGC 301
 70LysSerSerHisLeuValPhe.....SerGluAspLe 81
 302 ATCGTGTGGGGCGCGGACCATCGCTCTTACACTGAGCCGCCCGT 351
 81 uPhePheIleTyrLeuLeuProProIleIlePheAsnAlaGlyPheGlnV 98
 352 CTCTCTCTTCTACCTGTGCTGCCCCCATCGTGTGGACCGCGCTACTCA 401
 98 aLysLysLysGlnPhePheArgAsnPheValThrIleMetLeuPheGly 114
 402 TGCCCAACCGCTCTTCTCGGCAACCTGGGACCATCTCTGTTCAGGCC 451
 115 AlaValGlyThrIleLeuSerCysThrIleIleSerLeuGlyValThrG 131
 452 GTCTGGGTACCGGTGTGAACCGCGCCACACCGGGCTGTCCCTACGG 501
 131 nPhePheLysLysLeuAspIleGlyThrPheAspLeuGly.....AspT 146
 502 CGTCTTCCTCAGCGGCTCATGGGCGACCTGCAGATTGGGCTGGTACT 551
 146 yrLeuAlaIleGlyAlaIlePheAlaAlaThrAspSerValCysThrLeu 162
 552 TCCTCCTTTTGGAGCCTCATGGCGGTGTGGACCGCGGCGCTCCTG 601
 163 GlnValLeuAsnGln...AspGluThrProLeuValPheValThrLeuVal 178
 602 GCCGTGTTGAGGAGTCCATGTCAACGAGGCTCTGTTCATCATCGTCT 651
 178 eGlyGluGlyValValAsnAspAlaThrSerValValValPheAsnAlaI 195

652 CGGGAGTCGCTGTGAACGACGAGTCACCGTGTCTGTACAATGTGT 701
 195 leGlnSerPheAspLeuThrHisLeuAsnHisGluAlaAlaPheHisLeu 211
 702 TTGAATCTTTTCTGGCGCTGGAGGTGACAAGCTGACGTGGCGTGGACTGC 751
 212 LeuGlyAsnPheLeuTyrLeuPheLeuSerThrLeuLeuGlyAlaAl 228
 752 GTGAGGGCATAGTCTCTCTCTCTGGTGAGC...CTGGGGGGGACCGCT 798
 228 aThrGlyLeuIleSerAlaTyrValIleLysLysLeu...TyrPheGlyA 244
 799 GGTGGGGTGGTCTTCTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 848
 244 rHisSerThrAspArgGluValAlaLeuMetMetLeuMetAlaTyrLeu 260
 849 AGCATGTGGGTATCATCGAGCGCGCTTCTGTCTCTCTCTCTCTCTCTCT 898
 261 SerTyrMetLeuAlaGluLeuPheAspLeuSerGlyIleLeuThrValPh 277
 899 TCCTACCTGACGTCGAGATGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 948
 277 ePheCysGlyIleValMetSerHisTyrThrTrpHisAsnValThrGluS 294
 949 CTTCTGTGCATCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 998
 294 erSerArgIleThrThrLysHisThrPheAlaThrLeuSerPheLeuAla 310
 999 AGTCGGCCACCGCGTGGCTACACCATGAGATGCTGTCTGTCTGTCTGTCT 1048
 311 GluThrPheIlePheLeuTyrValGlyMetAspAlaLeuAspIleAspLy 327
 1049 GAGACCATCATCTTCATGTCTGTGGTATCTCGGCGTGAACCCGCTTCAT 1098
 327 strArgSerValSerAspThrProGlyThrSerIleAlaValSerSerI 344
 1099 CTGG.....ACCTGGAACACGCGCTTCTGTCTGTCTGTCTGTCTGTCT 1133
 344 leLeuMetGlyLeuValMetValGlyArgAlaAlaPheValPheProLeu 360
 1134 TGGTC.....TTCATCTCCGHTACCGGCGCATCGGTGTGTCTGTCTGT 1177
 361 SerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGluLysIleAsnPh 377
 1178 ACCTGCTTCTGAACCGCTACCGCATGTGGTGTGGAGCCCATTTGAC... 1225
 377 eAsnMetGlnValValIleThrTrpSerGlyLeuMetArgGlyAlaValS 394
 1226CAGGTGGTCTCTCTACGGGGCGCTG...CGCGGGGCGGTGG 1265
 394 erMetAlaLeuAlaTyrAsnLysPheThrArgAlaGlyHisThrAspVal 410
 1266 CTTTCCCTCTGGTG.....GTCTTCTGTGTGGAGACAGAGTCT 1303
 411 ArgGlyAsnAlaIleMetIleThrSerThrIleThrValCysLeuPheSe 427
 1304 AAGGAGAAGAACCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1353
 427 rThrValValPheGlyMetLeuThrLysProLeuIleSerTyrLeuLeup 444
 1354 CGTCATCTTCCAGGGCTCATCAAGCCTCTGTGGTGTGGTGTGGTGTGGT 1403
 444 roHisGlnAsnAlaThrThrSerMetLeuSerAspAspAsnThrProLys 460
 1404 TGAAGAGG.....AGCAGACCGCGGAGACCTCGG 1432
 461SerIleHisIleProLeuLeuAspGlnAspSerPheIleG 474
 1433 CTCACGAGAAGGTGCACGCGCGCTTTCGAC..... 1465
 474 uProSerGlyAsnHisAsnValProArgProAspSerIleArgGlyPheL 491

1466CACATCCTCTCGGCCATCGAGGACATATCCGGA..... 1498
 491 euThrArgProThrArgThrValHisTyrTyr.....TrpArg 503
 1499CAGATCGGGCACAATTATCTCAGACACAAGTGGTCC 1534
 504 GlnPheAspAspSerPheMetArgProVal 513
 1535 CACTTCGACAGGAAGTCTCTCAGCAGGGTTC 1564

seq_name: /SIDS6/gcgdata/geneseq/geneseq/NAL1993.DAT:051524

seq_documentation_block:

ID Q51524 standard; DNA; 4151 BP.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

1466CACATCCTCTCGGCCATCGAGGACATATCCGGA..... 1498
 491 euThrArgProThrArgThrValHisTyrTyr.....TrpArg 503
 1499CAGATCGGGCACAATTATCTCAGACACAAGTGGTCC 1534
 504 GlnPheAspAspSerPheMetArgProVal 513
 1535 CACTTCGACAGGAAGTCTCTCAGCAGGGTTC 1564

seq_name: /SIDS6/gcgdata/geneseq/geneseq/NAL1993.DAT:051524

seq_documentation_block:

ID Q51524 standard; DNA; 4151 BP.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

723 GATCCACCACATCTCCAGCATGCTCCCGGAGAGCTGCCTGCTGATTGGG 772
 55 leGlyLeuGlyThrGlyValThrIleLeuLeuIleSerLysGlyLysSer 71
 773 TGGGGCTGCTGTGGGGGCTGATC.....AAGGGCGTGGGC 810
 72 SerHisLeuValPheSerGluAspLeuPhePheIleTyrLeuLeu 88
 811 GAGAAGCGCCCTCTCTGTCAGTCGGAGGTCTTTTTCCTCTCTCTGCTGCC 860
 88 opIleIlePheAsnAlaGlyPheGlnValLysLysLysGlnPhePheA 105
 861 TCCCATCATCTGGACGGCGGTACTTCTGCGCGCTGCGGCGAGTTCACCG 910
 105 rgAsnPheValThrIleMetLeuPheGlyAlaValGlyThrIle..... 119
 911 AGAACCTGGGCACCATCTCTATCTTCGCGGTGGGCGACGCTGTGGAAC 960
 120IleSerCysThrIleIleSerLeuGlyValTh 130
 961 GCCTTCTTCTGGGGGCTCATGTACGCGGTGTGCTGCTGGCGGCGGA 1010
 130 rGlnPhePheLysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrL 147
 1011 GCAG.....ATCAACAACATCGGGCTCTGGACAAC 1042
 147 euAlaIleGlyAlaIlePheAlaAlaThrAspSerValCysThrLeuGln 163
 1043 TGCTGTGGCAGCATCATCTGGCGGTGGACCGCTGGCGCTGCTGCC 1092
 164 ValLeuAsnGlnAspGlu.....ThrProLeuLeuTyrSerLeuValPheG 179
 1093 GTCTTCGAGGAATCCACATCAATGAGCTGTGCTGCACATCTCTGGTCTCGG 1142
 179 yGluGlyValValAsnAspAlaThrSerValValValPheAsnAlaIleG 196
 1143 CGAGTCCCTGCTCAACGACGCGCTCACCTGTGTGTGTATCACCTCTTTG 1192
 196 loSerPheAspLeuThrHisLeuAsnHisGluAlaAlaPheHisLeuLeu 212
 1193 AGGAGTTT.....GCCAACTACGACACCGGGGCTGCTGGACATCTGC 1236
 213 GlyAsnPheLeuTyrLeuPheLeuSer.....ThrLeuLeuG 226
 1237 CTGGGCTTCTGAGCTTCTCTGCTGGCGCTGGCGGGGTGTTCTGTTGG 1286
 226 yAlaAlaThrGlyLeuIleSerAlaTyrValIleLysLysLeuTyrPheG 243
 1287 CGTGGTCTACGGGGTATCGAGCCTTCACTCCCGA.....TTCA 1327
 243 lyArgHisSerThrAspArgGluValAlaLeuMetMetLeuMetAlaTyr 259
 1328 CGCCCCACATCGCGGTATCGAGCGCTCTTCTGCTCTCTCTACAGCTAC 1377
 260 LeuSerTyrMetLeuAlaGluLeuPheAspLeuSerGlyIleLeuThrVa 276
 1378 ATGGCTTACTGTGGCGCGAGCTCTTCCACCTGTCTGGGCGAGTTCGAC 1427
 276 lPhePheCysGlyIleValMetSerHisTyrThrTrpHisAsnValThrG 293
 1428 CATCGCTCGGGAGTGTGATCGGCGCTTACGTGGAGGCGAAGCATCTCGC 1477
 293 luSerSerArgIleThrThrLysHisThrPheAlaThrLeuSerPheLeu 309
 1478 ACAAGTCGCACACCACTCAAGTACTCTCTGAAGATGTGGAGCAGCGTC 1527
 310 AlaGluThrPheIlePheLeuTyrValGlyMetAspAlaLeu...AspI 325
 1528 AGCAGACCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1577
 325 eAspLysTyrArgSerValSerAspThrProGlyThrSerIleAlaValS 342

Transgenic mice having this gene are useful as animal model for the analysis of the effect of sodium ion/hydrogen ion anti-transport carrier gene activation on the regulation of blood pressure. The transgenic animals are useful as models of hypertension and hypercardia.

Sequence 4151 BP; 754 A; 1429 C; 1174 G; 794 T; 0 other;

alignment_scores:

Quality: 441.00

Ratio: 1.353

Percent Similarity: 56.499

Percent Identity: 25.303

alignment_block:

US-09-271-584A-2 x Q51524

Align seg 1/1 to: Q51524 from: 1 to: 4151

23 LeuAsnLeuPheValAlaLeuLeuCysAlaCysIleValLeuGlyHisLe 39

673 ATCTCCCTCTGGATCTGTGGCTGCTCATGACATAGGTTTCCAGT 722

39 uLeuGluGlu...AsnArgTrpMetAsnGluSerIleThrAlaLeuLeu 55

810 GAACCTGGAGTGTGCTCTCTTTGTAAGTCTTTCCCGTATTTTTCGTAAGTAG 859
259
860 ACAAACAACACTCCTCCTCTCTCTCGTATTTATGACAACACTTCTTCC 909
260
910 CCCTAATGATTCTGGTATTCTCTAAGATACATGCTGCAGAGGTGT 959
269 AspLeuSerGlyIleLeuThrValPhePheCysGlyIleVal..... 282
|||||.....
960 GGTCTCTCCGCATTTGTTCTATACCTCTTCACAGAAATGTAATCGCGA 1009
282
1010 GTCATTGTAGCTTTTACATCTTAGTGTGATGTTAATATCTTGGAAAGACAT 1059
283Met..SerHisTyrThrT 288
|||.....
1060 ATTTAGGCTGCTAATATAGTGTACTGTAGTTATGAAGCGCTACACTT 1109
288 rpHisAsnValThrGluSerSerArgIleThrThrLysHisThrPheAla 304
: |||.....
1110 TCTCAAAATCTCTCAGAAAGTTCACAGAGTTTCGTATCTCTTTTTCAC 1159
305 ThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrValGlyMetAs 321
: |||.....
1160 TTGATATCTTCGGTAGCAGAAATCTTACCGTTTATACATGGGATTTGA 1209
321 p...AlaLeuAspIleAspLysTrp..... 328
|||.....
1210 TATTGCCATGAGCAGCATAGCTGGTCCCATGTTGGGTTTATCCTTTCT 1259
328
1260 CTATTGTATCCTCATTTACTGATGTCAGTGAATGTATGCAGTGTAGTC 1309
329ArgSerValSerAs 333
|||.....
1310 AGTGTGTAAATCCTTGCATTACCTTTGCTTCTGCTTCAAGACTGA 1359
333 pThr.....Prog 336
|||.....
1360 CATCAGTGTATTATTGGCTGGCTAGCTGACTAAATGCTTTTATCCTG 1409
336 lyThrSerIle..AlaValSer.....SerIleLe. 345
|||.....
1410 GCTGATCGCTTCATATACCATGATGTTTTCGATTCGGATTTACCTATATG 1459
345 uMetGlyLeuValMetValGlyArgAlaAlaPheValPheProLeuSerP 362
: |||.....
1460 TTCGCAATGCTTTCACCGCGGCTGTCATGTATTGGGTGTCAT 1509
362 heLeuSerAsnLeuAlaLysLysAsnGlnSerGluLysIleAsnPheAsn 378
: |||.....
1510 ATTTGGTCAACCTATTAGACAG...GAGAACCAAGATACCTATGAAG 1556
379 MetGlnValValIleTyrTrpSerGlyLeuMetArgGlyAlaValSerMe 395
|||.....
1557 CACCAAAAGCCCTTTGTGTATGTAGTACTT...CGAGGGCAATGGCAAT 1603
395 talaleuAlaTyrAsnLysPhe.....ThrArgAlaGlyHisThrAspV 410
: |||.....
1604 TGCATTGCACTTCAATCCTATCATGATACCAAGAGGTGCAC..... 1646
410 alaArgGlyAsnAlaIleMetIleThrSerThrIleThrValCysLeuPhe 426
: |||.....
1647GGCCAAATCATCTTTTACTGCAACCAACATATTGTTGTTGTC 1688
427 SerThrValValPheGlyMetLeuThrLysProLeuIleSerTyrLeuLe 443
: |||.....
1689 ACGGTTTTTACTAATAGGAGGTTCACAGGTAAATGTTGGAAGCTTTC.. 1736

443 uproHisGlnAsnAlaThrThrSerMetLeuSerAspAsnThrProL 460
: |||.....
1737GAAGTTGTAGTGACGAT..... 1754
460 ysSerIleHisIleProLeuLeuAspGlnAspSerPhe.....Ile 473
: |||.....
1755CTTGAT...GACTCCATGCTCGAAGGCTTT 1781
474 GluProSerGlyAsnHisAsnValProArgProAspSerIleArgGlyP 490
: |||.....
1782 GAAGACAGGATCATCATGATGCTCCTCTCTTTTAGCATTTGAGCTTC 1831
490 eleuThrArgProThr.....A 496
: |||.....
1832 ATCTCAGCAGGATACATCATCAGGAAGCAGGTTCAAGATGAAGCTGA 1881
496 rgThrValHisTyrTyrTrpArgGlnPhe.....AspAspSerPhe 509
: |||.....
1882 AGGAGTTTTCACAAAACCACTACATCATTCACCGCGTTGGACAAAACCTT 1931
510 MetArgProValPhe 514
1932 CTGACTCCGTTCTTC 1946

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:222593

seq_documentation_block:

ID 222593 standard; cdna; 714 BP.

XX

AC

XX

XX

DT

XX

XX

DE

XX

KW

KW

KW

XX

OS

XX

PN

XX

PD

XX

XX

PF

XX

PR

XX

PR

XX

PA

XX

PA

XX

PA

XX

PI

XX

DR

XX

DR

XX

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

: ::::: |||::: |||::: |||
 4604 ATTATTAGGCGGTCTAATTGGTTAACT..... 4575
 69 yLysSerSerHisLeuLeuValPheSerGluAspLeuPhePheIleTyrl 86
 ||| ::::: ::::: |||::: ::::: |||::: |||::: |||::: |||::: |||:::
 4574 .AAAATGGGCAGAAATGAATTAATTCAGCTCAGAGTGTTTTTTAGTAATGA 4526
 86 euLeuProProIleIlePheAsnAlaGlyPheClnValLysLysLysGln 102
 ::::: |||::: |||::: ||| ::::: ||| ::::: ||| ::::: |||
 4575 TTATTGCCACGCCTTATTTAGAGAAGGGAGCTAATCAATCGAAGACA 4476
 103 PhePheArgAsnPheValThrIleMetLeuPheGlyAlaValGlyThrIl 119
 ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
 4475 ATGATGAAAAAATTCGGCATGATCTCTTTCTGGCCCTTTTGTTGGTGTGT 4426
 119 eileSerCystThrIleIleSerLeuGlyValThrGlnPhePheLysLysL 136
 :||| ::::: |||::: ||| |||
 4425 AATT.....ACCITTAGTGGAGTCGGCTGGACC.....CTGCATA 4391
 136 euASPilleGlyThrPheAspLeuGlyAspTyrLeuAlaIleGlyAlaIle 152
 : ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
 4390 TCAITTTGSCCTACTTTACCATTTAGCTGCTTGTCTGCTTTTGAGCGGCT 4341
 153 pheAlaIlaThrAspSerValCysThrLeuClnValLeuAsnGlnAspGl 169
 :||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
 4340 TTAGGGCCAACGAGTGGCGGTGAGGTTCGTTATFACAGTTCGAATTCGA 4291
 169 uthrPro...LeuLeuTyrSerLeuValPhePheGlyGluGlyValValAsnA 185
 :||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
 4290 CATTCTCGAACGTCGATGAACATTTTATCTGGTGAAGGACTCATTAAATG 4241
 185 sPaIaThrServalValPheAsnAlaIleGlnSerPheAspLeuThr 201
 :||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
 4240 ATGCATCAGGCGTTACAGCCVTTCAAITTCGATTGATCGC 4203
 202 HisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTyrLe 218
 :||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
 4202 GCTTTA.....TTACGGGGTCATTTTCGCGCGT 4174
 218 u.....PheLeuLeuSerThrLeuLeuGlyAlaAlaThrG 230
 :||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
 4173 GGAAGCGGCAGGACCTTGGTTATTTCCAGCATTTGGTGGAGCGGTAGTCG 4124
 230 lyLeuIleSerAlaTyrValIleLysLysLeuTyrPheGlyArgHisSer 246
 :||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
 4123 GAATCGTGTAGTCTATTAAACAGCAAAATGTT.....CATTTA 4083
 247 ThrAspArgGluValala.....LeuMetMe 255
 :||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
 4082 CTTGAAACTCGATGCGCAAAAGATGTTACAGGCTACTTGTGTTGATGAAT 4033
 255 tLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPheAspLeuSerG 272
 :||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
 4032 ACTCTTGGCATTTATVTGCTTATTTAATTTACAGATGATTGGCGCTTCTG 3983
 272 lyIleLeuThrValPhePheCysGlyIleValMetSer..... 284
 :||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
 3982 GAATTATTCGTCGTGATGCTGCAATTAATCAAGCGCTCFAGTTTGAA 3933
 285HistTyrThrTrpHisasnValThrGluSerSerAr 296
 :||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
 3932 AAAAGTTTCGCTTTTGAAGCTGAGCTGCAAAATGTTGGCGAAGT..... 3868
 296 gileThrThrLysHisThrPheAlaThrLeuSerPheLeuAlaGluThrP 313
 :||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
 3887ATTGGAGTTTGATTGAGTTCACITTAATGCCT 3854
 313 heilePheLeuTyrValGlyMetaspAlaLeuAspIleAspLys..... 327
 :||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
 3853 TAGTTTCTTATTCTTGTAGCAATTAACCTTCACAGTGTTCGTCGCAAT 3804
 328 TrpArgSerValSerAspThrProGlyThrSerIleAlaValSerIl 344
 :||| ::::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::


```
XX Oryza sativa.
XX OS
XX PN WO9947679-A2.
XX PD
XX PF 23-SEP-1999.
XX PP 18-MAR-1999; 99WO-CA00219.
XX PR 18-MAR-1998; 98US-0078474.
XX PS 15-JAN-1999; 99US-0116111.
XX PA (BLUM/) BLUMWALD E.
XX PA (APSE/) APSE M.
XX PA (SNED/) SNEDDEN W.
XX PA (AHAR/) AHARON G.
XX PI Blumwald E, Apse M, Snedden W, Aharon G;
XX PP WPI; 1999-571840/48.
XX PR
XX PT Nucleic acid molecules encoding sodium/proton transport polypeptides,
XX PS useful in genetic engineering salt tolerance in crop plants
XX PP Disclosure; Fig 8D; 93pp; English.
XX CC The invention relates to an isolated nucleic acid molecule encoding
XX CC a plant Na/H antiporter (PNHX) transporter polypeptide, or a fragment
XX CC and capable of increasing salt tolerance in a cell. This sequence
XX CC corresponds to an Na/H antiporter transporter gene fragment from the
XX CC rice expressed sequence tag (EST) library clone C91861. The Na/H
XX CC transporter polypeptides provide a means of intracellular salt
XX CC management, particularly in plants. The sequences are useful for
XX CC producing transgenic plants that are capable of surviving in soil
XX CC with high salt levels that would normally inhibit growth of the crop
XX CC species. This would be useful in farming land in areas that are generally
XX CC considered unproductive through salt accumulation and poor irrigation,
XX CC e.g. in India, Australia, and prairies in USA or Canada. Commercial
XX CC crops, such as potatoes, tomatoes, brassica, cotton, sunflower,
XX CC strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley,
XX CC atriplex, sorghum, alfalfa, salicornia and others would benefit from
XX CC increased salt tolerance.
XX SQ Sequence 268 BP; 43 A; 55 C; 105 G; 61 T; 4 other;

alignment_scores:
    Quality: 190.00      Length: 66
    Ratio: 3.276        Gaps: 1
    Percent Similarity: 87.879    Percent Identity: 57.576

alignment_block:
US-09-271-584A-2 x 222607 ..
Align seg 1/1 to: 222607 from: 1 to: 268

15 SerAspHisAlaSerValValAlaLeuAsnLeuPheValAlaLeuLeuCy 31
   ::::::::::::::::::::|::::::::::::::::::|
72 GCGGACTACGGCTCGATCGCGCGGTGGGATGTTCTGTCGCTCATCTG 121
   ::::::::::::::::::::|::::::::::::::::::|
31 salacystleValLeuGlyHisLeuLeuGluAsnArgTrpMetAsnG 48
   ::::::::::::::::::::|::::::::::::::::::|
122 CGTCGTCATCGTCGTCGCGACCTCTCGAGGAGCCGATGGATGAACG 171
   ::::::::::::::::::::|::::::::::::::::::|
48 luSerIleThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeu 64
   ::::::::::::::::::::|::::::::::::::::::|
172 AGTCATCATCCCGCGTATCATCTCGGTGGGTACTGGAGGAGTGNTT 221
   ::::::::::::::::::::|::::::::::::::::::|
65 LeuIleSerLysGlyLysSerSerHisLeuLeuValPheSerGlu 79
   ::::::::::::::::::::|::::::::::::::::::|
222 GNATGGTGTGCGAGTGGTGAAGACACTCGGNATACTGTTGTTCACCGAG 267
   ::::::::::::::::::::|::::::::::::::::::|

seq_name: /SID56/gcdata/geneseq/geneseqn/NA1998.DAT.V52272
```

```
seq_documentation_block:
ID V52272 standard; DNA; 15363 BP.
XX AC
XX V52272;
XX DT 23-OCT-1998 (first entry)
XX DE Streptococcus pneumoniae genome fragment SEQ ID NO:139.
XX KW Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
XX KW computer readable medium; vaccine; pharmaceutical composition; ds.
XX OS Streptococcus pneumoniae.
XX PN WO9818931-A2.
XX PD 07-MAY-1998.
XX PF 30-OCT-1997; 97WO-US19588.
XX PR 31-OCT-1996; 96US-0029960.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;
XX PI Kunsch CA, Rosep CA;
XX DR WPI; 1998-272225/24.
XX PT Computer-readable medium with recorded Streptococcus pneumoniae
XX PT polynucleotide sequences - useful in diagnostic kits and assays, and
XX PT pharmaceutical compositions and vaccines for Streptococcus
XX PT pneumoniae
XX PS Claim 1; Page 931-939; 1409pp; English.
XX CC The present invention describes a computer readable medium which has
XX CC the nucleotide sequences SEQ ID NO:1 to 391 (V52134 to V52524) recorded
XX CC on it, or a representative fragment or a sequence at least 95% identical
XX CC to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in SEQ ID NO:1
XX CC to 391 (V52134 to V52524) are genomic fragments from Streptococcus
XX CC pneumoniae. The present invention also describes an isolated nucleic acid
XX CC molecule encoding a homologue of any of the fragments of the S.pneumoniae
XX CC genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced
XX CC by a process comprising: (a) screening a genomic DNA library using as a
XX CC probe a target sequence defined by any of the sequences in SEQ ID NO:1
XX CC to 391, identifying members of the library which contain sequences
XX CC that hybridise to the target sequence and isolating the nucleic acid
XX CC molecules from the members; or (b) isolating mRNA, DNA or cDNA produced
XX CC from an organism, amplifying nucleic acid molecules whose nucleotide
XX CC sequence is homologous to amplification primers derived from the
XX CC fragment of the S. pneumoniae genome to prime the amplification and
XX CC isolating the amplified sequences. The computer readable medium can be
XX CC used in a computer-based system for identifying fragments of the
XX CC S. pneumoniae genome of commercial importance, or expression modulating
XX CC fragments of the S. pneumoniae genome. Products from the present
XX CC invention can be used in diagnosis kits and assays, and pharmaceutical
XX CC compositions and vaccines for S. pneumoniae.
XX SQ Sequence 15363 BP; 4357 A; 3360 C; 2915 G; 4731 T; 0 other;

alignment_scores:
    Quality: 185.50      Length: 486
    Ratio: 0.733        Gaps: 21
    Percent Similarity: 52.058    Percent Identity: 20.576

alignment_block:
US-09-271-584A-2 x V52272/rev ..
Align seg 1/1 to reverse of: V52272 from: 1 to: 15363
```

```

1585 |||::: ::::: |||:::||||::: ::: :|||
ACCGACGGTCTGGCATACAGTGACCTTTATGCTCAACGGTTCTGTCTT 1536

315 eLeutyrvValGlyMetAspAlaLeuAspIleAspLysTrpArgSerValS 332
|||:::||||:::
1535 TCTGATTATTAGGATGGAGTTG.....GAAATGATG 1504

332 exAspThrProGlyThrSerIleAlaValSerIleLeuMetGlyLeu 348
::::: |||::: ::: :|||::: |||
1503 CAGAACCATTCTTGACCATTCCAATCTATAATCCTACTTTTATTCCTA 1454

349 ValMetValGlyArgAlaAlaPheValPheProLeuSerPheLeuSer.. 364
::::: ::: :|||::: ::: |||:::
1453 TCTCTTATGCCTTACCTTGTGCTCTTTTGTGCAATTCGTTTTATTATGAT 1404

365 .....AsnLeuAlaLysLysAsnGlnSerGluLysIleAsnP 377
::::: |||::: ::: :|||::: |||
1403 CTATGCTATTATGCTATPAGAACCCGACGCCCTCAAGAAAAGCTAAATA 1354

377 heAsnMetGln.....ValValIleTrpSerGlyLeuMetArgGly 391
|||::: ::::: ::|::|::: |||:::
1353 AGTATATGAAGACATGTTCTCTTGACCTTTCAGGT...GTTTAGGGA 1307

392 AlaValSerMetAlaLeuAlaTrpAsnLysPheThrArgAlaGlyHisTh 408
::|::|::|::|::|
1306 ACGGTGTCGATTGCT..... 1292

408 rAspValArgGlyAsnAlaIleMetIleThrSerThrIleThrValCysL 425
|||::: |||::: |||:::
1291 .....ACGATTCTCTTGATACCAAGTAATCTAGAACAG...G 1258

425 euPheSerThrValValPheGlyMetLeuThrLysProLeuIleSerTy 441
::: :|||::: |||::: |||::: |||::: |||:::
1257 AGTATCCTCTTTTGTCTTTTCTTGTGTGAGGTGTGACGCTGTGTACTTT 1208

442 Leu.....LeuProHisGlnAsnAlaThrThrSerMetLe 453
||| |||::|::|
1207 TTACACAGGTCTCTTGTGCTGCTCAT.....CT 1179

453 uSerAspAspAsnThrProLysSer.....lleHisIleProLeuL 467
|||::|::|::: ::::: ::|::|::: |||
1178 TTCTGTGAAGAGGAGGAAGCAAGGATTATCTTATGCATATCGCCATTT 1129

467 euAspGln 469
||::: ||:::
1128 TGAATGAA 1121

seq_name: /SIBS6/gcgdata/geneseq/geneseqn/NA1997.DAT:v74657
seq_documentation_block:
ID: v74657 standard; DNA; 1294 BP.
XX
XX v74657;
XX
XX
XX 16-MAR-1999 (first entry)
DE Staphylococcus aureus contig SEQ ID #346.
XX
XX Computer readable medium; vaccine; S.aureus infection; immu-
KW cellulitis; eyelid infection; food poisoning; osteomyelitis
KW skin infection; surgical wound infection; scalded skin synd-
KW toxic shock syndrome; ds.
XX
OS Staphylococcus aureus.
XX
XX EF786519-A2.
XX
XX 30-JUL-1997.
PD
XX
XX 07-JAN-1997; 97EP-0100117.
PF
XX
XX 05-JAN-1996; 96US-0009861.
PR
XX
```

PR 05-JAN-

53 enI:enI]eG]vI:enG]vThrG]vVa]ThrI]aI:


```

386 rGlyLeuMetArgGlyAlaValSerMet.....AlaLeuAlaTyrAsnL 401
|||:||||: |||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
7392 AGGNGTCATCGCGAGTATATTGACTTTTACTTATGACATTTACATGATAA 7441
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
401 ysPheThrArgAlaGlyHisThrAspValArgGly..... 412
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7442 AAGAAACATTCCTGGGAATTAACAATATAGAAAAATTTAAACGTAACAA 7491
413 .....AsnAlaIleMetIleThrSerTh 420
||||:||||: |||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
7492 ATACATGAACCATGGCTATTAGTTTACACGCTGTGATTTTAAATGTTACT 7541
420 rIleThrValCysLeuPheSerThrValValPheGly 432
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7542 CATTCAGTTATCTCTTCCTTCCAAACGTTTITGGC 7578

seq_name: /cqn2.6/ptodata/1/ina/5A_COMB.seq:US-08-920-827-6

seq_documentation_block:
; Sequence 6, Application US/08920827
; Patent No. 5770375
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,827
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORGANISM: Staphylococcus epidermidis
; STRAIN: Clinical isolate SE-22
; US-08-920-827-6

```

```

alignment_scores:
  Quality: 132.00      Length: 546
  Ratio: 0.557        Caps: 25
  Percent Similarity: 43.407  Percent Identity: 20.879

```

```

alignment_block:
US-09-271-584A-2 x US-08-920-827-6
Align seg 1/1 to: US-08-920-827-6 from: 1 to: 8654

1 MetLeuAspSerLeuValSerLysLeuProSerLeuSerThrSerAspHi 17
6107 ATGGCGACTAATTTATTAGTCATGCTCATAGTTTATTTCACCTCTGAGTCA 6156
17 sLaSerValValAlaLeuAsnLeuPheValAlaLeuLeuLeuLeuLeu 34
6157 TCGTCAACTAAGAAAGGTTCGGGCTATGTTGCATTAATAGTCTCTATTG 6206
34 leValLeuGlyHisLeuLeuGluGlu.....AsnArg 44
6207 TGACATCTACATATTTTATTATGAAAAATACCAGATGTCATTCGAAATAAG 6256
45 .....TrpMet.....AsnGluSerIleTh 51
6257 TTTATTGCTGTTTCGATTACCATGGATGCTTCAATTGATTAATTTAGA 6306
51 tAlaLeuLeuIleGlyLeuGlyThr.....GlyValThrIleLeuLeu 66
6307 TTTAAGATTAGATGGTTTAAGTTTAATGTTTCGGCTTAATTTATTCGGCTAA 6356
66 leSerLysGly.....LysSerSerHisLeuLeuValPheSerGlu 79
6357 PGGTGTGGGTGTTATTTTATGCTACGCAATATTATCCACAGTACG 6406
80 Asp.....LeuPheIleIleTyrLeuLeuProIleIleIlePheAs 93
6407 GACAATCTCTCTAGATTTTTCATCTATTACTA..... 6439
93 nAlaGlyPheGlnValLysLysGlnPhePheArgAsnPheValThr 110
6440 .....T 6440
110 leMetLeuPheGlyAlaValGlyThrIleIleSerCysThrIleIleSer 126
6441 TATTATGTTTCAGTATGATGGCATTGTAATAGCTAATAATACCATCTTA 6490
127 LeuGlyValThrGlnPhePheLysLysLeuAspIleGlyThrPheAspLe 143
6491 ATGTATGTA.....TTTGGGAACCTCACAAGTATTTCCTCATTCTGCT 6534
143 uGlyAspTyr..... 146
6535 TATATCTTATGTTGTAACAATAATGGTGAAGTCAATTAGCGCCATTCAT 6584
147 .....LeuAlaIleGlyAlaIlePheAlaAlaThrAspSerVal 159
6585 CTTTCATGATTACAGTGTGTTGGTGGCTAGCGTTATTAAACAGGATTTATC 6634
160 CysThrLeuGlnValLeuAsnGlnAspGluThrProLeuLeuTyrSerLe 176
6635 .....ATTTTATAT...AT 6645
176 uValPheGlyGluGlyValValAsnAspAlaThrSerValValValPheA 193
6646 CATTACAGGAACAACAACAAATTACTGAT.....ATCTTTAATCA 6683
193 snAlaIleGlnSerPheAspLeuThrHisLeu..... 203
6684 ACGCAATGCAATTTTCACGACATCCTTTATTATACCAATGATTTTGTATGC 6733
204 AsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTyrLeuPhe... 219
6734 TATTATTAGTGCTTTTACCAAATCTGCACAAATTTCCGTTTCATATTGG 6783
220 Leu.....LeuSerThrLeuLeuGlyAl 227
6784 TTACCAAGGCCATGGCAGCACCTACACCAGTAAGTCTTATCTTCATTC 6833

```

227 aAlaThr..... 229
 6834 GCGCAACAAATGGTAAAGCGCTGGAATCTTTTACTATTAGATTACACCTT 6883
 230GlyLeuLeuSerAlaTyrValIleLysLysLeuTyrPheGlyArg 244
 6884 TATGGGACCTTAGTAATGTTATATATTATACAGTCACATTTGTGGT... 6930
 245 HisSerThrAspArgGluValAlaLeuMetMetLeuMetAlaTyrLeuSe 261
 6931CTAATACTATGTTATTGGCAT 6953
 261 rTyrMetLeuAlaGluLeuPheAspLeuSerGlyIleLeu.....T 275
 6954 TTTAACTGCTTTACGACAATACGAGCTAAAAGGTATACTCGCTTATTC 7003
 275 hrValPhePheCysGlyIleValMetSer.....His 285
 7004 CAATAAGTCAATPAGGTATGATATGACAATGGTAGTCTAGGTGGCGGT 7053
 286 TyrThrTrpHisAsnValThrGluSerSerArgIleThrThrLysHisIst 302
 7054 TATGCTCACCACACATCAGATGAATTGCTPAAGTTTATATTATTAGTTT 7103
 302 rPheAlaThrLeuSerPheLeuAlaGluThrPheIlePhe.....L 316
 7104 ATTTGCTGCTTATTCCTATTTAATGAATCATCGCGGTTTAAATGTGCAT 7153
 316 euTyrValGlyMetAspAlaLeuAsp..... 324
 7154 TATTTATGGCGTGGTATCATTCATGATCAGAGTCGGAACACGTGATATT 7203
 325IleAspLysTrpArgSerValSerAspThrProGlyThrSerII 339
 7204 CGTTTGCCTAAATGGTATCGCTAAAGTCTCC.....CCTAAAT 7241
 339 eAlaValSerSerIleLeuMetGlyLeuValMetValGly.....A 353
 7242 GCATATTGCTATGTTGCTCGCTGCATTATCATGGCAGGTGCTCTTTT 7291
 353 tgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAlaLysLys 369
 7292 TAATGGCTTTTAAAGTAAGGAATGTTTTAGATTGCTTAACTAAAGCA 7341
 370 AsnGlnSerGluLysIleAsnPheAsnMetGlnValValIleThrIrpSe 386
 7342 AACGAACCTTGATCAATATGGCTCGTATTAACGTTTGTAFTATTCAAT 7391
 386 rGlyLeuMetArgGlyAlaValSerMet.....AlaLeuAlaTyrAsnL 401
 7392 AGGTGTCATCCGAGTATATGACTTTTACTTATGCATTTACATGATAA 7441
 401 ysPheThrArgAlaGlyHisThrAspValArgGly..... 412
 7442 AAGAAACATCTCGGGAAATTTACAAATAGAAAATTTAAACGTAACAA 7491
 413AsnAlaIleMetIleThrSerTh 420
 7492 ATACATGAACCATGGCTATTATTAGTTTACCAGCTGTGATTTTATGTACT 7541
 420 rIleThrValCysLeuPheSerThrValValPheGly 432
 7542 CATCCCAAGTATCTTCCTTGTGTCACCAACGTTTTTGGC 7578

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-921-177-6

seq_documentation_block:

; Sequence 6, Application US/08921177

; Patent No. 5798211

GENERAL INFORMATION:

; APPLICANT: Ohno, Tsuneya

; APPLICANT: Matsuhisa, Akio

APPLICANT: Uehara, Hirotsugu
 APPLICANT: Eda, Soji
 TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 6300 Sears Tower, 233 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/921,177
 FILING DATE: 29-AUG-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/362,577
 FILING DATE: 27-MAR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Rin-Laures, Li-Hsien
 REGISTRATION NUMBER: 33,547
 REFERENCE/DOCKET NUMBER: 19036/32420
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/474-6300
 TELEFAX: 312/474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 8654 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE:
 ORGANISM: Staphylococcus epidermidis
 STRAIN: Clinical isolate SE-22
 US-08-921-177-6

alignment_scores:

Quality:	132.00	Length:	546
Ratio:	0.557	Gaps:	25
Percent Similarity:	43.407	Percent Identity:	20.879

alignment_block:

US-09-271-584A-2 x US-08-921-177-6

Align seg 1/1 to: US-08-921-177-6 from: 1 to: 8654

```

1 MetLeuAspSerLeuValSerLysLeuProSerLeuSerThrSerAspHis 17
6107 ATCGGGCACTAAATTTATTAGTCATGCTCATAGTCTTTATTACTCTGAGTCA 6156

17 sAlaSerValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34
1 : ::::: ::::: ::::: ||||| ::::: ::::: ::::: :::::
6157 TCGTCAACTAAGAAGGTGCGGCCTATGTTGCATTAATAAGTCCTCATTC 6206

34 leuValLeuGlyHisLeuLeuGlu.....AsnArg 44
::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
6207 TGACATCTACATATTTTATTATGAATAACCAGATGTGATTCGAAATAAG 6256

45 .....TrpMet.....AsnGluSerIleTh 51
||||| ::::: ::::: ::::: :::::
6257 TTATTGCTGTGATTCACATGAGTGCCITCAATTCATTAATTATTAGA 6306

51 rAlaLeuLeuLeuGlyLeuGlyThr.....GlyValThrIleLeuLeu 66
||| ||||| ::::: ::::: ::::: :::::
6307 TTTAAGATTACATGGTTTAACTTTAATGTTCGGCTTAATTTATTTTCGTAA 6356
```


TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 8654 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Staphylococcus epidermidis
STRAIN: Clinical Isolate SE-22
US-08-362-577C-6

alignment_scores:		
Quality:	132.00	Length: 546
Ratio:	0.557	Gaps: 25
Percent Similarity:	43.407	Percent Identity: 20.879

```
alignment_block:
US-09-271-584A-2 x US-08-362-577C-6
..
```

Align seg 1/1 to: US-08-362-577C-6 from: 1 to: 8654

1 MetLeuAspSerLeuValSerLysLeuProSerLeuSerThrSerAspHi 17

6107 ATGGCGACTAATTTATTAGTCATGCTCATAGTTTATTCACTCTGAGTCA 6156

17 sAlaSerValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34

6157 TCGTCAACTAAGAAAGGTTGCGGGCTATGTTGCATTAAATAGCTCCTATTG 6206

34 leValLeuGlyHisLeuLeuGluGlu.....AspArg 44

6207 TGACATCTACATATTTTATTATGAAAAATACCAGATGTGATTCGAAATAG 6256

```

45 .....TroMet.....AsnGluSer]eTh 51

```

6257 TTTATTGCTGTTCCGATTACCATGGGATGCCCTCAATTGATATTATTTACA 6306

51 rA]aLeu]eu]eG]vLeuG]vThr- C]vVa]ThrT]a]oU]oU] 66

[illegible]

ALCOHOL 99

6357

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

[illegible][illegible]

ATTENTION: NEW YORK TIMES MAGAZINE

0450 1.....

```

119 # -----
120 # -----
121 # -----

```

6444 TATTATGCTCCGCTGTGATGGCCATGGTATACCACTCTTA 6490

[illegible]

04971 AIGIATGTA.....TTTGGGAACCTACAAGTATTTCCCTCATTTCTTGGCT 6534

143	ugly	aspyr	146
		...		

6535 TATAATCTATTTGGTACAAATAATGTGAAAGTCAATTAGCGCCATTCAAT 6584

147LeuAlaIleGlyAlaIlePheAlaAlaThrAspServal 159

6585 CTTTCATGATTACAGTGTTTGGTGGCTAGCGTTATTAAACAGGATTATC 6634

[illegible]


```

seq_documentation_block:
: Sequence 6, Application US/08920827
: Patent No. 5770375
: GENERAL INFORMATION:
: APPLICANT: Ohno, Tsuneya
: APPLICANT: Matsuhisa, Akio
: APPLICANT: Uehara, Hirotsugu
: APPLICANT: Eda, Soji
: TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
: NUMBER OF SEQUENCES: 25
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 6300 Sears Tower, 233 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606-6402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/920,827
: FILING DATE: 29-AUG-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/362,577
: FILING DATE: 27-MAR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Rin-Laures, Li-Hsien
: REGISTRATION NUMBER: 33,547
: REFERENCE/DOCKET NUMBER: 19036/32420
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312/474-6300
: TELEFAX: 312/474-0448
: TELEX: 25-3856
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 8654 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: Genomic DNA
: ORIGINAL SOURCE:
: ORGANISM: Staphylococcus epidermidis
: STRAIN: Clinical Isolate SE-22
: US-08-920-827-6

alignment_scores:
Quality: 122.00 Length: 695
Ratio: 0.391 Gaps: 36
Percent Similarity: 44.892 Percent Identity: 18.849

alignment_block:
US-09-271-584a-2 x US-08-920-827-6/rev ..
Align seg 1/1 to reverse of: US-08-920-827-6 from: 1 to: 8654
14 ThrSerAspHisAlaSerValValAlaLeuAsnLeuPheValAlaLeuLe 30
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4665 ACCTCGGAACATTCACAACTT.....ATTCTATATCT 4634
30 uCysAlaCysIleValLeuGlyHisLeuLeuGluGluAsnArgTrpMetA 47
|||||
4633 TTGCTTA.....TGCTCA 4620
47 snGluSerIleThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIle 63
|||||
4619 ATCATTTTGAACAACATGATTGATGATG..... 4588
64 LeuLeuIleSerLysGlyLysSerHisLeuLeuValPheSerGluAs 80

```

```

:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4587 ATAATCATATTAAATGAGAAAGTTGGTTTTTA.....TC 4553
80 pLeuPhePheIleTyrLeuLeuProPheIle.....P 92
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4552 ATTGTTTTTCTATATCTTTTAGGAATTTTGGTGGCATGTACTTCCTA 4503
92 heAsnAlaGlyPheGlnValLysLysGlnPhePheArg....AsnPhe 107
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4502 TTCAAACTTCAATCAACTCAAGCTAAGCCAAATTTACCCGTTTCATCTTT 4453
108 ValThrIleMetLeuPheGlyAlaValGlyThrIleIleSerCysThrI 124
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4452 TATGCTTCAACCATATCTTTTGCTGTCGGAACAATA.....TCCTTACT 4409
124 eile.....SerLeuGlyValThrGlnPhePheL 134
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4408 TGTGCTTAATATATCTTATTCACCAAGTATTAAACACCAAGATCTTTT 4359
134 yLysLeuAspIleGly...ThrPheAspLeuGlyAspTyrLeuAlaIle 149
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4358 CTAAGCAAAACACTTAACATATATCTTGGGTCCTAGGAGGACTATTAGGT 4309
150 .....GlyAlaIlePh 153
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4308 ATCTATTTAACCTGGAACCTTATTATTATTCACCAAGATTAGCGCAGCACT 4259
153 eAlaAlaThrAspSerValCysThrLeuGlnValLeuAsnGln..... 167
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4258 TACTGTTGTTATTACGGTT...ACAGGGCAAAATTTATTATGGGTGTAATCA 4212
168 ..AspGluThrProLeuLeu.....TyrSerLeuVal 177
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4211 TTGATACATTTTGGATTATTGGCGCCCATCAACAATCTTTTACATTTT 4162
178 PheGlyGluGlyValValAsnAspAlaThrSerValValValPheAsnAl 194
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4161 AAGGTGTTGGAATTTATTTTAACTTACGGAATTTATTTTATGAATTA 4112
194 aileGlnSerPheAspLeuThr...HisLeuAsnHisGluAlaAlaPheH 210
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4111 TGTTCGTAGACATCTCTGTTAATACACATAAGAATACACCAATAGTATCT 4062
210 isLeuLeuGlyAsnPheLeuTyrLeuPhe..... 219
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4061 GGTGCTTATTGGATTCGTATTGG..TTTCGTCCGCCAATCCAAACAACA 4013
220 .....LeuLeuSerThrLe 224
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
4012 ATCAACAGCACACTGGCTCAACACACACACTCATCATATTTCCTTCACT 3963
224 uLeuGlyAlaAlaThrGlyLeuIleSerAlaTyrValIleLysLysLeu 241
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3962 CATATCATTTTAGCGTGTGTACAAATAGCGTTATTATTATTA...ACATTAG 3916
241 yrPheGlyArg.....HisSerThrAspArgGluValAlaLeu 253
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3915 TATTTAATCGTAGTCTGAAAAATATCATCTCATATAAAACACTAGGTAAG 3866
254 MetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPheAspLe 270
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3865 ATTAATCTATTATTATTATTGAGGGGACTAGGAATGGCATTTCGTAAC 3816
270 uSerGlyIleLeuThrValPhePhe..... 278
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3815 ATCTAATATTATTATTATGCTTCTTAGGAGCAGCTTTAACTACTATAA 3766
279 .....CysGlyIleValMetSerHisTyrThr 287
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3765 TTGCTATGATGGCCAAATGATTATGGGTATTATTATTCGATCATTTCCGT 3716
288 TrpHisAsnValThrGluSerSerArgIleThrThrLysHisThrPheAl 304
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

3715 TTGTTGGGGT...TCACCTAAAAACAAAATAAATACATCAGT...AAGTTAGG 3676

304 aThrLeuSerPheLeuAlaGluThrPheIlePheLeu..... 316

3671 TGGATTATTATGTATAGCTATAGGTATTATTTTATTATACGATTATTTAAA 3622

317TyrValGlyMetAspAla... 322

3621 TTCCTTATTCTGAATAAAAAATTATCCGCTTATATTAATGTAGATTCTCTCA 3572

323LeuAspIleAspLys.....Tr 328

3571 TACAAATAAATGCTATTTCGAAGCTGCCAACAAAGATTTTGCTCTTGTGTG 3522

328 pArg..SerValSerAspThrProGlyThrSerIleAlaValSerSerIle 344

3521 GCAGTTTATGAGTGATCTCTTATAGTAAACATTTATGTAGAAATCATTA 3472

344 eLeuMetGlyLeuValMetValGlyArgAlaIlePheValPhePro.... 359

3471 TATTTTAAAAATTA.....AAGATAGCCTCTATTTTTCATATTT 3434

360LeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGluLys 374

3433 ATTAATTTGATTACTTATTATTTTGAATAAAACAATATACAAACCAA 3384

375 IleAsnPheAsnMetGlnValIleTrp..... 384

3383 AGTATGCTTATTAAATTAATATATTTCATTTTAAATAATAGTGTATA 3334

385TrpSerGlyLeuMetArgGlyAlaValSerMetAlaLeuAla 399

3333 ATAACCTCACAAATAATATGAGG..... 3309

399 yTrAsnLysPheThrArgAlaGlyHisThrAspValArgGlyAsnAlaIle 415

3308TGAACATCCATGACCAAGCAAGCATC 3282

416 MetIleThrSer.....ThrIleThrValCysLeuPhe 427

3281 GTTATTCTCTATTAATCTATTCATTCTCATCTCACACTTGTAGTTTAG 3232

427 rThrValValPheGlyMetLeuThrLys..... 436

3231 TTACATTACTTTCTCTAAATTAATCTCTAAAAACAGAAGATTAAAAATA 3182

436 436

3181 TAAAAACTATAAAGCAAGCCCAATTTACTCTCTTCAACAACCTGGTATA 3132

437ProLeuIleSerTyrLeuLeu.....ProHisGlnAsnAl 448

3131 CAAGCCCTCTTAACAGCTATATTTCATCTACACTCCCACTTTAGGACA 3082

448 aThrThrSerMetLeuSerAspAsnThrProLysSerIleHisIleP 465

3081 ACTCCAAAAACTTTTAGTAAAAAATCATAGCAAGATATCTAAAAACACT 3032

465 rOleuLeu.....AspGlnAspSerPheIleGluPro... 475

3031 CASTGTGTAATATTACAAATTATGAAAAGACTATTATATGAGGAACCT 2982

476SerGlyAsnHisAsnValProArgProAspSerIleAr 488

2981 CAAATACACTTAGTAATGAAACGTAAACAACAAAACCTTTAAATCT... 2937

488 gGlyPheLeuThrArgProThrArgThrValHisTyrTyrTrpArgGlnP 505

2936AAAATGATTACATATTAT... 2916

505 hAspAspSerPheMetArg..... 511

2915TCATTTCTTATCAGCTAAAAATTTACAAATGATTCAAACTAAA 2874

```

512 .....ProValPheGlyArgGlyph 519
2873 TCCTCAATAAATCCACTTATTTCACCAATAATGAAGTATTTTCATA 2824
519 eValPropheValProGlySerProThrGluArg 530
2823 TATTAAATCTATCCTCTCAAACCAACGAAAAA 2790

seq_name: /cgn2_6/ptodata/1/lna/5A_COMB.seq;us-08-921-177-6

seq_documentation_block:
; Sequence 6, Application US/08921177
; Patent No. 5798211
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/921,177
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus epidermidis
; STRAIN: Clinical Isolate SE-22
US-08-921-177-6

alignment_scores:
Quality: 122.00 Length: 695
Ratio: 0.391 Gaps: 36
Percent Similarity: 44.892 Percent Identity: 18.849

alignment_block:
US-09-271-584A-2 x US-08-921-177-6/rev ...

Align seg 1/1 to reverse of: US-08-921-177-6 from: 1 to: 8654
14 ThrSerAspHisAlaSerValAlaLeuAsnLeuPheValAlaLeuLe 30
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

4665 ACCTCGGAACATTCACAACTT.....ATTCTCTACT 4634
30 uCysAlaCysIleValLeuGlyHisLeuLeuGluAsnArgTrpMetA 47
|||||
4633 TTGCTTA.....TGGCTCA 4620
47 sngluSerIleThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIle 63
|||||
4619 ATCATTTTAGAAACATGATTGAATGCTATG..... 4588
64 LeuLeuIleSerLysGlyLysSerSerHisLeuLeuValPheSerGluAs 80
|||||
4587 ATAATCATATTAAATGAGAAATGGTGTTTTA.....TC 4553
80 pleuPheIleTyrLeuLeuProProlIle.....P 92
:|||||
4552 ATGTGTTTTCTATATCTTTTAGGAATTTTGTGGCAATGTACTTCCTA 4503
92 heAsnAlaGlyPheGlnValLysLysGlnPhePheArg...AsnPhe 107
:|||||
4502 TTCAAACTTCAATCAACTCAAGGCTAAGCCAAATTTACCCGTTTCATCTTT 4453
108 ValThrIleMetLeuPheGlyAlaValGlyThrIleIleSerCysThrIle 124
:|||||
4452 TATGCTTCAACCATATCTTTTCTGCTCGGAACAATA.....TGCTTACT 4409
124 eile.....SerLeuGlyValThrGlnPhePheL 134
:|||||
4408 TGTGCTTAAATATCATATTATCCACAAGTATTAAACACAGAATTCCTTTT 4359
134 ysLeuLeuAspIleGly...ThrPheAspLeuGlyAspTyrLeuAlaIle 149
:|||||
4358 CTAAGCAAACTTAACATATATCTGGGTCCTAGGAGACTATTAGGTGTT 4309
150GlyAlaIlePh 153
4308 ATCTATTAACTGGAACCTTATTATTATTAACAGATTAGGCGCAGCAT 4259
153 eaAlaIleThrAspSerValCysThrLeuGlnValLeuAsnGln..... 167
4258 TACTGTGTATTACGGTT...ACAGGCAAAATTATTATGGGTGTAATCA 4212
168 ..AspGluThrProLeuLeu.....TyrSerLeuVal 177
|||||
4211 TTGATACATTTTGATATTGGGCGCCCATCAACATCTTTTACATATT 4162
178 PheGlyGluGlyValValAsnAspAlaThrSerValValPheAsnAl 194
|||||
4161 AAAGGTGTGGAATTATATTTTAAATTACTGGAATTATATTATGAATTA 4112
194 atIleGlnSerPheAspLeuThr...HisLeuAsnHisGluAlaAlaPheH 210
:|||||
4111 TGTTCGTAGACATCTGTTAATAGACATAAGAATACACCAATAGTATTCT 4062
210 isLeuLeuGlyAsnPheLeuTyrLeuPhe..... 219
|||||
4061 GGTTCCTATTGGATTCTGATTATGG..TTTCGTCGCCCAATCAACAACA 4013
220LeuLeuSerThrIle 224
4012 ATCAACAGCACACTGGCTCAACACACACACTCATCATATTTTGCCTCACT 3963
224 uLeuGlyAlaAlaThrGlyLeuIleSerAlaTyrValIleLysLeuT 241
:|||||
3962 CATATCATTTTACGCTGGTACAAATAGCGTTATTATATTA...ACATAG 3916
241 yrPheGlyArg.....HisSerThrAspArgGluValAlaLeu 253
:|||||
3915 TATTTAATCGTAGCTGAAATATCATCTACTCATATAAAACACTAGGTAAG 3866
254 MetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPheAspLe 270
:|||||
3865 ATTAAATCTATTATTATTGAGGGGACTAGGAATGGCATTTCGTAAC 3816

270 uSerGlyIleLeuThrValPhePhe..... 278
|||||
3815 ATCTAATATTATTAAATGCCTTCTTAGGAGCAGCTTTAACTACTATATA 3766
279CysGlyIleValMetSerHisTyrThr 287
|||||
3765 TTGCTATGATGGCCAAATGATTATGGGTATTATTATCGATCATCTCGGT 3716
288 TrpHisAsnValThrGluSerSerArgIleThrThrLysHisThrPheAl 304
:|||||
3715 TTGTTGGGT...TCACCTAAAAACAAATAACATCACGT...AAGTAGG 3672
304 aThrLeuSerPheLeuAlaGluThrPheIlePheLeu..... 316
:|||||
3671 TGGATTATTATGTATAGCTATAGGTATTATTATTACGATTATTTTAA 3622
317TyrValGlyMetAspAla... 322
:|||||
3621 TTCCTTATTCTGAATAAAAAATTATCGCTTATATTAAATGTAGATCTCTCA 3572
323LeuAspIleAspLys.....Tr 328
:|||||
3571 TACAAATAAATGTCTATTTCGAACCTGCCAAACAAAGATTTTGTCTTGTG 3522
328 pArg...SerValSerAspThrProGlyThrSerIleAlaValSerSerIl 344
:|||||
3521 GCAGTTTATGAGTGTATCTTCTTATAGTAACATTTATGTAGATCATA 3472
344 eLeuMetGlyLeuValMetValGlyArgAlaAlaPheValPhePro.... 359
:|||||
3471 TATTTTAAATTA.....AAGATAGCCTTCTATTTTCAATTT 3434
360LeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGluLys 374
:|||||
3433 ATTAAATGTATTACTTTTGTGAAATAAAACAATATACAAACCAA 3384
375 IleAsnPheAsnMetGlnValIleTrp..... 384
|||||
3383 AGTAATGCTTATTAAATAATATATTTCATTAAATAATAGTGTATA 3334
385TrpSerGlyLeuMetArgGlyAlaValSerMetAlaLeuAlaT 399
:|||||
3333 ATAACCTCACAAATAATTATGAGG..... 3309
399 yAsnLysPheThrArgAlaGlyHisThrAspValArgGlyAsnAlaIle 415
:|||||
3308TGAACATCCATGCCAACAGCACGCATC 3282
416 MetIleThrSer.....ThrIleThrValCysLeuPheSe 427
:|||||
3281 GTTATTCTTCTTACTTATTTCATTTCATTCCTCACAACTTGTAGTTTAG 3232
427 rThrValValPheGlyMetLeuThrLys..... 436
:|||||
3231 TTACATTACTTCTCTAAATTAATCTCTAAACACAGAGATATTAATAA 3182
436 436
3181 TAAAACTATAAAAGCAAGCCCAAAATTTACTCTTTTCAACACTGGTATA 3132
437ProLeuIleSerTyrLeuLeu.....ProHisGlnAsnAl 448
|||||
3131 CAAGCCCTCTTAACAGCTATATTCTTCTACTACACTCCCACCTTTAGGACA 3082
448 aThrThrSerMetLeuSerAspAsnThrProLysSerIleHisIleP 465
:|||||
3081 ACTCAAAAACCTTTTAGTAAAAAATCATAGCAAAAGTATCTAAAAACACTC 3032
465 rOleuLeu.....AspGlnAspSerPheIleGluPro... 475
:|||||
3031 CAGTGTTAGAATATTACAATTTATGAAAAGATCATTTATATGAGGAACCTC 2982

```
476 .....SerGlyAsnHisAsnValProArgProAspSerIleAr 488
|||||.....
2981 CAAATAACACTTAGTAATAAGCGTAACAACAAACATTTTAAATTC 2937
|||||.....
488 gGlyPheLeuThrArgProThrArgThrValHisTyrTyrTrpArgGlnP 505
|||||.....
2936 .....AAAAATGATTACATTTATAT 2916
|||||.....
505 heAspAspSerPheMetArg 511
|||||.....
2915 .....TCATTCCTTAATCAGCTAAATAATACAAATGATTCACAAATAA 2874
|||||.....
512 .....ProValPheGlyGlyArgGlyPh 519
|||||.....
2873 TCCTCAATTAATCCACTTTATTTTACCAATTAATGGAAGTATTCATA 2824
|||||.....
519 evalProPheValProGlySerProThrGluArg 530
|||||.....
2823 TATTAATTCATCCCTTCAAAACCAACGAAAG 2790
|||||.....
```

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-362-577C-6

```
seq_documentation_block:
; Sequence 6, Application US/08362577C
; Patent No. 5807673
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362.577C
; FILING DATE: 27-MAR-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus epidermidis
; STRAIN: Clinical Isolate SE-22
US-08-362-577C-6
```

alignment_scores:
Quality: 122.00 Length: 695
Ratio: 0.391 Gaps: 36
Percent Similarity: 44.892 Percent Identity: 18.849

```
alignment_block:
US-09-271-584A-2 x US-08-362-577C-6/rev ..
Align seg 1/1 to reverse of: US-08-362-577C-6 from: 1 to: 8654

14 ThrSerAspHisAlaSerValValAlaLeuAsnLeuPheValAlaLeuLe 30
|||||.....
4665 ACCTCGGACATTCACACTT.....ATTCTATATCT 4634
|||||.....
30 uCysAlaCysIleValLeuGlyHisLeuGluGluAsnArgTrpMetA 47
|||||.....
4633 TTGCTTA.....TGGCTCA 4620
|||||.....
47 snGluSerIleThrAlaLeuLeuLeuGlyThrGlyValThrIle 63
|||||.....
4619 ATCATTTTAGAACATGATTTATGAATGCTATG..... 4588
|||||.....
64 LeuLeuIleSerLysGlyLysSerHisLeuLeuValPheSerGluAs 80
|||||.....
4587 ATAATCATATTAATGAGAAAGTTGGTTTAA.....TC 4553
|||||.....
80 pLeuPhePheIleTyrLeuLeuProIleIle.....P 92
|||||.....
4552 ATTTGTTTTTCTATATCTTTTAGGAATTTTGTGGCATGTTACTTCT 4503
|||||.....
92 heAsnAlaGlyPheGlnValLysLysLysGlnPhePheArg....AsnPhe 107
|||||.....
4502 TTCAAACTTCAATCAACTCAAGGCTAAGCCAATTTACCCGTTTCATCT 4453
|||||.....
108 ValThrIleMetLeuPheGlyValGlyThrIleIleSerCysThrI 124
|||||.....
4452 TATGCTTCAACCATATCTTTTGCCTCGAACATA.....TCCTTACT 4409
|||||.....
124 eIle.....SerLeuGlyValThrGlnPhePheL 134
|||||.....
4408 TGTGCTTAATATCATTTATTCACACAGATTAACACACAGATTTCTTT 4359
|||||.....
134 ysLysLeuAspIleGly...ThrPheAspLeuGlyAspTyrLeuAla 149
|||||.....
4358 CTAGCAAAACACTTAACATATCTTTGGGTCTTAGGAGGACTATTAGG 4309
|||||.....
150 .....GlyAlaIlePh 153
|||||.....
4308 ATCTATTTAACTGGAACTTATTATTATTTACCAAGATTAGCGCAGCA 4259
|||||.....
153 eAlaAlaThrAspSerValCysThrLeuGlnValLeuAsnGln..... 167
|||||.....
4258 TACTGTTGTTTACGGTT...ACAGGGCAAAATTTATTATGGGTGTAAT 4212
|||||.....
168 ..AspGluThrProLeuLeu.....TyrSerLeuVal 177
|||||.....
4211 TTGATACATTTGGATTATTGGGCGCCCATCAACAATCTTTTACATTAT 4162
|||||.....
178 PheGlyGluGlyValValAsnAspAlaThrSerValValValPheAsn 194
|||||.....
4161 AAAGGTGTGGAATTATTTTAACTTACTGGAATTTATTTATGAATTA 4112
|||||.....
194 alleGlnSerPheAspLeuThr...HisLeuAsnHisGluAlaAlaPheH 210
|||||.....
4111 TGTTCTGATAGACATCTCTGTTAATAGACATAAGAAATACACCAATAG 4062
|||||.....
210 isLeuLeuGlyAsnPheLeuTyrLeuPhe..... 219
|||||.....
4061 GGTTCCTTATTGATTGCTATTGTTGG.TTTCGTCGCGCAATCCAAACA 4013
|||||.....
220 .....LeuLeuSerThrLe 224
|||||.....
4012 ATCAACAGCACACTGGCTCAACACACACACTCATCCATATTTGCTTCA 3963
|||||.....
224 uLeuGlyAlaAlaThrGlyLeuIleSerAlaTyrValIleLysLysLeu 241
|||||.....
3962 CATATCATTTAGCGTTGGTACAATAGCGTTATTATATTA...ACATTAG 3916
|||||.....
```

```
241 yrPheGlyArg.....HisSerThrAspArgGluValAlaLeu 253
|||||
3915 TATTATTCGTAGTCGAAATATCATCTACTCATATAAAACACATAGGTAAG 3866
254 MetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPheAspLe 270
:::
3865 ATTAATCTATTATTTATGGAGGATAGTAGGAATGCATTCGTAAC 3816
:::
270 uSerGlyLeuThrValPhePhe..... 278
|||||
3815 ATCTAATATATTTAATGCTTCTCTAGGAGCAGCTTTAACTACTATAA 3766
279 .....CysGlyIleValMetSerHisTyrThr 287
|||||
3765 TTCTGATGATGGCCCAAAATGATGATGATGATGATGATGATGATGATGATG 3716
288 TrpHisAsnValThrGluSerSerArgIleThrLysHisThrPheAl 304
:::
3715 TTCTTGGGT...TCACCTAAACAAATAACATCACGT...AAGTTAGG 3672
304 aThrLeuSerPheLeuAlaGluThrPheIlePheLeu..... 316
|||
3671 TGGATTATTATGATAGCTAGGTATTATTATTACGATTATTATTAA 3622
317 .....TyrValGlyMetAspAla... 322
|||
3621 TTCTCTTATTCGTAATAAAATATTCGCTTATATTAATGATGATGATGATG 3572
323 .....LeuAspIleAspLys.....Tr 328
:::
3571 TACAATAAATGCTATTTCGAATGCCCAACAAGATTTCGTTGTTG 3522
328 pArg...SerValSerAspThrProGlyThrSerIleAlaValSerIle 344
:::
3521 GCAGTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3472
344 eLeuMetGlyLeuValMetValGlyArgAlaAlaPheValPhePro.... 359
:::
3471 TATTAAAAATTA.....AAGTAGCCTCTTATTTCATTTTCAATTT 3434
360 .....LeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGluLys 374
:::
3433 ATTAATGATGATTACITTTATTTTGAATTAACAATATACAAACCAA 3384
|||
375 IleAsnPheAsnMetGlnValValIleTrp..... 384
|||
3383 AGTAATGCTTATTAAAAATTAATATATTCATTTAAAAATAATAGTGATA 3334
385 .....TrpSerClyLeuMetArgGlyAlaValSerMetAlaLeuAlaT 399
:::
3333 ATACCTCAAAATAATATATGAGG..... 3309
399 yrAsnLysPheThrArgAlaGlyHisThrAspValArgGlyAsnAlaIle 415
|||
3308 .....TGAACATCCATGACCAAGCAACGCATC 3282
416 MetIleThrSer.....ThrIleThrValCysLeuPheSe 427
:::
3281 GTTATTTCTTCATTTACTATTTCATTTTCATTTTCATTTTCATTTTCATTT 3232
427 rThrValValPheGlyMetLeuThrLys..... 436
|||||
3231 TTACATTACTTTCTCTAAATAAATCTCTTAAACAGAGATATTTAAAAATA 3182
436 ..... 436
3181 TAAAAACTATAAAGCAAAAGCCCAAAATTTACTCTTTCAACAACCTGGTATA 3132
437 .....ProLeuIleSerTyrLeuLeu.....ProHisGlnAsnAl 448
|||||
3131 CAAGCCCTCTTAACAGCTATATTTTCTACTACACTCCCACTTTTAGGACA 3082
```

```
448 aThrThrSerMetLeuSerAspAsnThrProLysSerIleHisIleP 465
:::
3081 ACTCAAAACTTTTAGTAAAAATCATAGCAAGATATCTAAACACTC 3032
465 roLeuLeu.....AspGlnAspSerPheIleGluPro... 475
|||||
3031 CAGTCTTAGAATATTACAATATGAAAAAGATCATTTATATGAGGAAC 2982
476 .....SerGlyAsnHisAsnValProArgProAspSerIleAr 488
|||||
2981 CAAATACACTTAGTAATGAACGTAAACAAACAAATTTAAATCT... 2937
488 gGlyPheLeuThrArgProThrArgThrValHisTyrTyrTrpArgGlnP 505
:::
2936 .....AAAAATGATTTACATTTAT... 2916
505 heAspAspSerPheMetArg..... 511
|||||
2915 .....TCATTCCTTAATCAGCTAAATTTACAAATGATTTCAAACTAAA 2874
512 .....ProValPheGlyGlyArgGlyPh 519
|||||
2873 TCCTCAATTAATCCACTTTATTTTCACCAATTAATGGAAGTATTTCTATA 2824
519 eValProPheValProGlySerProThrGluArg 530
:::
2823 TATTAAATCTATCCCTTCAAAACCAACGAAAG 2790
seq_name: /cgn2_6/ptodata/1/ina/5b_comb.seq:US-08-920-828-6

seq documentation block:
; Sequence 6, Application US/08920828
; Patent No. 5853998
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,828
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8654 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
```

210 isLeuLeuGlyAsnPheLeuThrLeuPhe.....LeuLeuSerThrLe 224
 ||||| :|||:||||| :|||
 4061 GGTTCCTTATTGGATTGCTGATTGG..TTTCGTCCGCCAATCCAAACAACA 4013
 220LeuLeuSerThrLe 224
 :|||:|||||:|||||
 4012 ATCAACAGCACACTGGCTCAACACACACACTCATCCATATTTGCTTCAC 3963
 224 uLeuGlyAlaAlaThrGlyLeuLeuSerAlaThrValIleLysLysLeu 241
 :|||:|||||:|||||:|||||
 3962 CATATCATTTAGCGTTGGTACAAATAGCGTTATTTATATTA...ACATTAG 3916
 241 yrPheGlyArg.....HisSerThrAspArgGluValAlaLeu 253
 |||:||||| :|||:||||| :|||:|||||
 3915 TATTTAATCGTAGCTGAAATATCATCTACTATAAAACACATAGGTAA 3866
 254 MetMetLeuMetAlaThrLeuSerTyrMetLeuAlaGluLeuPheAspLe 270
 :|||:|||||:|||||:|||||:|||||
 3865 ATTAATCATTTATTTATTTATTTAGGAGGATAC TAGGAATGCATTCGTAA 3816
 270 uSerGlyIleLeuThrValPhePhe..... 278
 :|||:|||||:|||||:|||||
 3815 ATCTAATATTTATTTAATGCGCTTCTTAGGAGCAGCTTTAACTACATATA 3766
 279CysGlyIleValMetSerHisTyrThr 287
 :|||:|||||:|||||:|||||:|||||
 3765 TTGCTATCATGGGCCAATGATTATGGGTATTATTATCGATCATTTCCGT 3716
 288 TrpHisAsnValThrGluSerSerArgIleThrThrLysHisThrPheAl 304
 :|||:|||||:|||||:|||||:|||||
 3715 TTGTTGGGT...TCACCTATAAAACAAAATAACATCAGCT...AAGTTAG 3672
 304 aThrLeuSerPheLeuAlaGluThrPheIlePheLeu..... 316
 :|||:|||||:|||||:|||||:|||||
 3671 TGGATTATATGATAGCTATGATGATTATTTATACGATTATTTTAA 3622
 317TyrValGlyMetAspAla... 322
 :|||:|||||:|||||:|||||:|||||
 3621 TTCCTTATTCGATAAAAAATTATCCGCTTATATTATATGATTCCTCA 3572
 323LeuAspIleAspLys.....Tr 328
 :|||:|||||:|||||:|||||:|||||
 3571 TACAATAAATGCTATTTTGGAACTGCCAACAAAGATTTTGTCTTTGTT 3522
 328 pArg..SerValSerAspThrProGlyThrSerIleAlaValSerSerIl 344
 :|||:|||||:|||||:|||||:|||||
 3521 GCAGTTTATGATGATGATCTCTCTATAGTAACATTTATGAGAATCAT 3472
 344 eLeuMetGlyLeuValMetValGlyArgAlaAlaPheValPhePro.... 359
 :|||:|||||:|||||:|||||:|||||
 3471 TATTTTAAATTA.....AGATAGGCTCTATTTTTCATTT 3434
 360LeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGlyLys 374
 :|||:|||||:|||||:|||||:|||||
 3433 ATTAATATGATTACTTTATTTTGGAAATAAAACAATATACAAACCAA 3384
 375 IleAsnPheAsnMetGlnValValIleTrp..... 384
 :|||:|||||:|||||:|||||:|||||
 3383 AGTAAATGCTATTTTAAATTAATATTTTTCATTTAAATAATAGTGTATA 3334
 385TrpSerGlyLeuMetArgGlyAlaValSerMetAlaLeuAla 399
 :|||:|||||:|||||:|||||:|||||
 3333 ATAACCTCACAAATATTTATGAG..... 3309
 399 yrAsnLysPheThrArgAlaGlyHisThrAspValArgGlyAsnAlaIle 415
 :|||:|||||:|||||:|||||:|||||
 3308TGAACATCCATGATCCAAAGCAACGATC 3282
 416 MetIleThrSer.....ThrIleThrValCysLeuPhe 427
 :|||:|||||:|||||:|||||:|||||
 3291 GTTATTTCTCTATTCATTCATTTTCATTTCTCACAACTTGTAGTTTAG 3232
 427 rThrValValPheGlyMetLeuThrLys..... 436


```
1601 .....TACAACTCTTCCCCCAACACATGCT 1629
183 lasnaspAlaThrSerVal.....ValValPheAsnAlaIleGlnSerP 198
1630 GGAGCGTGATCCTCGGCGCAACATCACCTGTGTCAGAGATCCAG.... 1675
198 heAspLeuThrHisLeuAsnHisGluAlaAlaPheHisLeu..... 212
1676 .....GTGCCCTTGCTCTTGAGA 1693
212 ..... 212
1694 GCATCCCGGTTGGAAATCGGCAAGTGTATGCTGTGTGATAACCCCTGGAC 1743
213 ....GlyAsnPheLeuTyr...LeuPheLeuLeuSerThrLeu..... 224
1744 TGGAGGCATTTCCCTCATCGCTCTGTTTCATATCTTCACCTCTTATTGCT 1793
225 .....LeuGlyAlaAlaThrGlyLeuIleSerAlaTyrValIle 237
1794 TGATCGCGCAATGGATCCACCATGGGGATGTTAGCAGACACTCAGCAT 1843
238 .....LysLysLeuTyrPheGlyArgHisSerThrAspArgL 250
1844 GCCACACCTTTGACTCCATCTACTTGGCCTGTGCTGGCTTCAACAGCAC 1893
250 uValAla.....LeuMetMetLeuMetAlaTyrLeuS 261
1894 GTGGCATCATCGCTGTGGAGGAGTCTTACGTCATCATCCTGGCAGA 1943
261 erTyrMetLeuAla.....GluLeuPhe..... 268
1944 CTCACCTCTTCTGCTGTGCTCGCCCTGTTGGGCGCTACGTGGTGCT 1993
269 AspLeuSerGlyIleLeuThrValPhe..... 277
1994 GCCTGACCAAGCATATATCTGTGTGGATTACCAACCTGCACCTGGCC 2043
278 .PheCys 279
2044 CTTCTGC 2050
```

seq_name: /cgn2_6/ptodata/1/ina/6_COMB.seq:US-08-602-809-3

```
seq_documentation_block:
; Sequence 3, Application US/08602809
; Patent No. 6008012
; GENERAL INFORMATION:
; APPLICANT: BERGSMAN, DEREK
; APPLICANT: ELLIS, CATHERINE
; TITLE OF INVENTION: HUMAN SOMATOSTATIN-LIKE R
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: P.O. Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: US
; ZIP: 19482-0980
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,809
; FILING DATE: 13-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCW/US95/16472
; FILING DATE: 15-DEC-1995
```

```
; APPLICATION NUMBER: US 08/357,675
; FILING DATE: 16-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestia, Paul F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: P50277
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 601-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1080 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-602-809-3

alignment_scores:
Quality: 96.50 Length: 371
Ratio: 0.652 Gaps: 22
Percent Similarity: 39.892 Percent Identity: 22.642

alignment_block:
US-09-271-584A-2 x US-08-602-809-3 ..
Align seg 1/1 to: US-08-602-809-3 from: 1 to: 1080

171 ProLeuLeuTyrSerLeuValPheGlyGluGlyValValAsnAspAlaTh 187
172 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
94 CCTACCTCTACAGCATCATCTTTGGTGGGAATA..... 129
187 rSerValValValPheAsnAlaIleGlnSerPheAspLeuThrHisLeuA 204
129 ..... 129
204 snHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTyrLeuPheLeu 220
130 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
130 ..TTGGAAACAGCTTGGTGGTGCATCTCTC 156
221 LeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIleSerAlaTyrValI 237
157 ATT.....TACTTTTA 167
237 elysLysLeuTyrPheGlyArgHisSerThrAspArgGluValAlaLeuM 254
168 CATGAAGCTG.....AAGACTGTGCCACGC 193
254 etMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPheAspLeu 270
194 TCTTTCTT.....CTCAATCTCGCTTGGCT.....GACTTA 225
271 SerGlyIleLeuThrValPhePheCysGlyIleValMetSer...HisTy 286
226 TGCTTTTCTGCTGACTTTGCCCCCTGTGGCAGCTCTATACCGCATGAGTA 275
286 rThrThrHisAsnValThrGluSerSerArgIleThrThrLysHisThrP 303
276 CCGCTGGCCCTTCGGCAATCACCTATGTAGATCGCTTCG..... 315
303 heAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrValGly 319
316 ..GCCAGCGTGAGCTCAACCTCTACGCCAGCTGTGTCTCTT..... 354
320 MetAspAlaLeuAspIleAspLysTrpArgSerValSerAspThrProGl 336
355 CTCACGTGCTCTCAGCATCGACCGCTAC..... 381
336 yThrSerIleAlaValSerSerIleLeuMetGlyLeuValMetValGlyA 353
381 ..... 381
```

```

seq_name: /cgn2_6/ptodata/1/1na/PCTUS_COMB.seq: PCT-US95-16472-3
seq_documentation_block:
; Sequence 3, Application PC/TUS9516472
; GENERAL INFORMATION:
; APPLICANT: Bergsma, Derk J
; APPLICANT: Ellis, Catherine E
; TITLE OF INVENTION: Human Somatostatin Receptor
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation/Corporate
; ADDRESSEE: Intellectual Property
; STREET: P. O. Box 1539-UW2220
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16472

```

Mon Mar 26 11:04:17 2001

```
353 rgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAlaLysLys 369
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
382 ..CTGCCATCGTCCACCAATGAAG.....TCTCGCCTTCGCCGACG 423
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
370 AsnGlnSerGluLysIleAsnPheAsnMetGlnValValIleTrpTrpSe 386
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
424 ATGCTGGTGGCCAAAGTCACCTGC.....ATCATCATCTGG..... 459
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
386 rgLeuMetArgGlyAlaValSerMetAlaLeuAlaTyrAsnLysPheT 403
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
460 ....CTGATGGCTGGCTGGCCAGTTTCCACGCTGTCATCCACCGA.... 501
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
403 hrArgAlaGlyHisThrAspValArgGlyAsnAlaIleMetIleThrSer 419
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
502 .....AATGTATACTTCATCCAGAAC 522
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
420 Thr...IleThrValCysLeuPhe.....SerThrVa 429
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
523 ACCAATATCACAGTGTGGCGGTTTCATTATGAGTCTCGGAATTCGACGT 572
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
429 IValPheGlyMet...LeuThrLysProLeuIleSerTyrLeuLeuProH 445
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
573 CCCATAGGGCTGGCCCTACCAAGAAATATTCTGGGCTTCTGTCCCTT 622
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
445 IsGlnAsnAlaThrThrSer...MetLeuSerAspAsnThrProLys 460
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
623 TCCTTATCATCTCTCACCGCTATACCTTATTGSAAGCTCTAAAGAG 672
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
461 SerIleHisIleProLeuLeuAspGlnAspSerPheIleGluProSerGl 477
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
673 GCTTATGAAATT..... 684
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
477 yAsnHisAsnValProArgProAspSerIle.ArgGlyPheLeuThrArg 493
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
685 .CAAAAGAACAAACCAAGAACAGATGACATCTTTAGGATAATTA..... 727
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
494 ProThrArgThrValHisTyrTyrTrpArgGln...PheAspAspSerPh 509
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
728 .....TGGCGATTGTCCTTTTCTCTCTTT 753
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
509 eMetArgProValPheGlyArgGlyPheValProPheValProGlyLys 526
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
754 T.....CCTGGGT 761
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
526 erProThrGlu 529
    |||:|||| |||:|||| |||:|||| |||:|||| |||:||||
762 CCCCCACCAAA 772
```


OM of: US-09-271-584A-2 to: GenEmbl.* out_format : pfs
Date: Mar 26, 2001 10:35 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=frame+p2n_model -DEV=xlp
-O=/cgn2_1/USPto_spool/US09271584/runat_15032001_151301_5671/app_query.fasta_1.604
-DB=GenEmbl -QFMT=fastcap -SUFFIX=p2n.rge -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000
-CGAPOP=4.500 -CGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62
-TRANS=human40.cdi -LIST=45 -DCCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM_ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09271584_cgn1_1_5190 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPPY
-WAIT -THREADS=1

Search information block:
Query: US-09-271-584A-2
Query length: 538
Database: GenEmbl.*
Database sequences: 1118133
Database length: -1736092196
Search time (sec): 2605.960000

score_list:

Sequence	Strd Orig	Zscore	EScore	Len	Documentation
gb_p11:AF106324	+ 2755.00	3469.57	4.2e-185	1614	! AF106324 Arabidopsis thaliana
gb_p11:AF056190	+ 2755.00	3469.54	4.3e-185	1619	! AF056190 Arabidopsis thaliana
gb_p13:TW021B04	- 2282.50	2833.28	1.2e-149	90019	! AF007271 Arabidopsis thaliana
gb_p11:AB038492	+ 2168.50	2725.00	1.3e-143	2637	! AB038492 Atriplex gmelini AgNH
gb_p11:AB033989	+ 2041.00	2565.95	9.1e-135	2218	! AB033989 Ipomoea nil Purple m
gb_p11:AB021878	+ 2023.50	2543.46	1.6e-133	2313	! AB021878 Oryza sativa O5NH1 n
gb_p11:AC009465	+ 1924.00	2380.82	1.9e-124	92234	! AC009465 Arabidopsis thaliana
gb_p11:ATAC011623	+ 1492.50	1837.47	3.4e-94	86022	! AC011623 Arabidopsis thaliana
gb_p11:AC033990	+ 1266.50	1576.92	1.1e-79	7545	! AC033990 Ipomoea nil Purple ge
gb_p11:AB015479	- 1118.00	1365.83	6.4e-68	80675	! AB015479 Arabidopsis thaliana
gb_in2:AF142676	+ 615.50	767.44	1.4e-34	2405	! AF142676 Drosophila melanogast
gb_p13:AF279671	+ 586.00	747.76	1.7e-33	421	! AF279671 Mesembryanthemum crys
gb_p13:DB7743	+ 581.00	717.84	7.9e-32	4408	! DB7743 Human mRNA for KIAA0267
gb_p13:AF030409	+ 581.00	717.75	8.0e-32	4452	! AF030409 Homo sapiens sodium-H
gb_in2:AF199463	+ 568.00	706.74	3.3e-31	2604	! AF199463 Drosophila melanogast
gb_p13:SCD9461	+ 564.00	667.46	5.1e-29	78500	! AF199463 Drosophila melanogast
gb_p12:AF279670	+ 542.00	691.40	2.4e-30	459	! AF279670 Mesembryanthemum crys
gb_h1:AC017375	+ 530.00	625.24	1.1e-26	73515	! AC017375 Drosophila melanogast
gb_in1:AF003590	- 530.00	611.05	7.0e-26	301639	! AB003590 Drosophila melanogast
gb_h1:AC008318	- 513.00	588.47	1.3e-24	338007	! AC008318 Drosophila melanogast
gb_p13:SPAC15A10	+ 502.50	595.79	5.0e-25	43702	! 297208 S.pombe chromosome I c
gb_h1:AC016779	- 486.00	563.30	3.2e-23	139626	! AC016779 Oryza sativa clone
gb_h1:RABNAHEX	+ 475.00	588.92	1.2e-24	2748	! M87007 Oryctolagus cuniculus N
gb_ro:RATNHXIII	+ 467.00	572.51	9.9e-24	5153	! M85300 Rat sodium-hydrogen exc
gb_p13:AF111173	+ 466.50	575.04	7.1e-24	3761	! AF111173 Homo sapiens sodium/H
gb_ro:AF100172	+ 466.50	574.00	8.2e-24	4174	! AF100172 Rattus norvegicus sod
gb_pat1:AR035574	+ 462.00	578.23	4.7e-24	2574	! AC035574 Sequence 8 from pater
gb_p17:HSU28043	+ 462.00	573.14	9.1e-24	2584	! U28043 Human plasma membrane N
gb_om:OPONHE3R	+ 459.00	568.34	1.7e-23	2859	! L42522 Didelphis virginiana Na
gb_ro:RATMATSTOM	+ 455.50	560.74	4.5e-23	3925	! L11236 Rattus rattus mRNA sequ
gb_om:RABNH2A	+ 454.50	559.77	5.1e-23	3813	! L13733 Rabbt ethylisopropylam
gb_ro:RABNH2X	+ 454.50	559.44	5.2e-23	3941	! L11004 Rat sodium-hydrogen exc
gb_p13:AF073299	+ 452.50	560.68	4.5e-23	2711	! AF073299 Homo sapiens Na+/H+ e
gb_in2:AF235935	+ 451.00	555.90	8.3e-23	3612	! AF235935 Drosophila melanogast
gb_p11:AB023156	+ 451.00	551.61	1.4e-22	5537	! AB023156 Homo sapiens mRNA for
gb_p13:AB023156	+ 444.00	550.13	1.7e-22	2665	! S81591 SLCA2-Na+/H+ exchanger
gb_p13:AF141350	+ 443.00	549.36	1.9e-22	2538	! AF141350 Homo sapiens sodium/H
gb_p13:AF141351	+ 443.00	549.36	1.9e-22	2538	! AF141351 Homo sapiens sodium/H
gb_p13:AF141352	+ 443.00	549.36	1.9e-22	2538	! AF141352 Homo sapiens sodium/H
gb_p13:AF141353	+ 443.00	549.36	1.9e-22	2538	! AF141353 Homo sapiens sodium/H

gb_p13:AF141354	+ 443.00	549.36	1.9e-22	2538	! AF141354 Homo sapiens sodiu
gb_p13:AF141355	+ 443.00	549.36	1.9e-22	2538	! AF141355 Homo sapiens sodiu
gb_p13:AF141356	+ 443.00	549.36	1.9e-22	2538	! AF141356 Homo sapiens sodiu
gb_p13:AF141357	+ 443.00	549.36	1.9e-22	2538	! AF141357 Homo sapiens sodiu
gb_p13:AF141358	+ 443.00	549.36	1.9e-22	2538	! AF141358 Homo sapiens sodiu
seq_name: gb_p11:AF106324					
seq_documentation_block:					
LOCUS	AF106324	1614 bp	mRNA	PLN	03-MAR-1999
DEFINITION	Arabidopsis thaliana sodium proton exchanger Nhx1 mRNA, partial cds.				
ACCESSION	AF106324				
VERSION	AF106324.1	GI:4324596			
KEYWORDS					
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
REFERENCE					
AUTHORS	Gaxiola,R.A., Rao,R., Sherman,A., Grisafi,P., Alper,S.L. and Fink,G.R.				
TITLE	The Arabidopsis thaliana proton transporters, AtNhx1 and Avp1, can function in cation detoxification in yeast				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (4), 1480-1485 (1999)				
MEDLINE	99145575				
REFERENCE	2 (bases 1 to 1614)				
AUTHORS	Gaxiola,R.A., Rao,R., Sherman,A., Grisafi,P., Alper,S.L. and Fink,G.R.				
TITLE	Direct Submission				
JOURNAL	Submitted (12-NOV-1998) Whitehead, Nine Cambridge Center, Cambridge, MA 02142, USA				
FEATURES	Location/Qualifiers				
source	1..1614				
	/organism="Arabidopsis thaliana"				
	/cultivar="Columbia"				
	/db_xref="taxon:3702"				
CDS	1..>1614				
	/codon_start=1				
	/product="sodium proton exchanger Nhx1"				
	/protein_id="AA016946.1"				
	/db_xref="GI:432459"				
	/translation="MDSLSVSKPLSLSTSDHASVVALNLFVALLCACIVLHLEENR				
	WMNESITALLIGTGVTIILISKGKSSHLVSEDLFFIYLPPIIFNAGFOVKKKQ				
	FFNFVITMLFGAVGTIIISGLVTQFFKKLDIGTDFDGLDYLAIGAFAATDSVC				
	TLOVLNDETPGLISLVFGGVNDATSVVFNAIQSFDTLHNHAAFLHLLGNFLYL				
	FLLSLTLLGAATGLISAVIKKLYFGRHSSTOREVLAAMLAAYLSYMLAEALFDLGGILTY				
	FFCGIVMSHYTNWNTVETSSRTTKHTFATLSFAETFIIFYVGMADLIDKWRVSQDT				
	PCPTSIASVSSILMLGLVAVGRAFPVPLSFLSNLAKKQSEKINFMQVVIWWSGLMRGA				
	VSMALATNKRTRAGTIDVRGNAMITSTIIVCLFSTVVFQMLTKPLISLYLLPQNATT				
	SMUSDONTPKSIHPLDDQDSFTPEGSHNVPRPDSIRGELTRTPRTVTHVYWRQFDDDS				
	FMRPVFGRGFVFPVFGSPTEPNPPDLISA"				
BASE COUNT	388 a 346 c 358 g 522 t				
ORIGIN					
alignment_scores:					
Quality:	2755.00	Length:	538		
Ratio:	5.121	Gaps:	0		
Percent Similarity:	100.000	Percent Identity:	100.000		
alignment_block:					
US-09-271-584A-2 x AF106324					
Align seg 1/1 to: AF106324 from: 1 to: 1614					
1 MetLeuAspSerLeuValSerLysLeuProSerLeuSerThrSerAspHis 17					
1 ATGTTGGATTCCTCTAGTCGAAACTGCGTTCGTTATCGACATCTGATCA 50					
17 salaSerValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34					

51 .CGCTTCTGCTGGTGGTGAATCTCTTTGTTGCACCTCTTTGTTGCTTTGTA 100
34 leValLeuGlyHisLeuLeuGluGluAsnArgTrpMetAsnGluSerIle 50
101 TTGTTCTTGGTCATCTTTTGAAGAGATAGATGGATGAAGATCCATC 150
51 ThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeuLeuIleSe 67
151 ACCGCTTGTGATTGGCTAGGCACCTGGTGTACCATTTTGTGATTAG 200
67 rLysGlyLysSerSerHisLeuLeuValPheSerGluAspLeuPheI 84
201 TAAAGGAAAAGCTCGCATCTTCGCTTTTGTAGTGAAGATCTTTTCTTCA 250
84 leTyrLeuLeuProPheIlePheAsnAlaGlyPheGlnValLysLys 100
251 TATATCTTTTCCACCCCATATATCAATGCGAGGTTTCAAGTAAAAAG 300
101 LysGlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaValGI 117
301 AAGCAGTTTTCGCAATTTCTGCTGACTATTATGCTTTTGGTGTGG 350
117 yThrIleIleSerCysThrIleIleSerLeuGlyValThrGlnPhePheL 134
351 GACTATTATTTCTGCAATCATATCTCTAGGTGTAACACAGTTCTTTA 400
134 ysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleGly 150
401 AGAGTTGACATTTGGAACCTTTGACTGGTGGTGATATCTTGTCTATTGGT 450
151 AlaIlePheAlaIaThrAspSerValCysThrLeuGlnValLeuAsnGI 167
451 GCCATATTTTGTGCAACAGATTGATGTACACTGCGAGGTTCTGAATCA 500
167 nAspCluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValVala 184
501 AGACGAGACCTTTGCTTTTACACTTGTATTCGAGAGGTTCTTGTGA 550
184 snAspAlaThrSerValValPheAsnAlaIleGlnSerPheAspLeu 200
551 ATGATGCAACCTGACTGTTGCTCTTCAACGCGATTTCAGAGCTTTGATCTC 600
201 ThrHisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTy 217
601 ACTCACCTAAACACGAGAGCTGCTTTTCATCTTCTTGAAACTTCTTGTA 650
217 rLeuPheLeuLeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIleSerA 234
651 TTTGTTCTCCTTAAGTACCTTGTGCTGGTGTGCAACCGGCTGATAAGTG 700
234 laTyrValIleLysLysLeuTyrPheGlyArgHisSerThrAspArgGlu 250
701 CGTATGTTATCAAGNAGTATATCTTTGGAAGGCACTCACTGACCGAGAG 750
251 ValAlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLe 267
751 GTTGGCCCTTATGATGCTTATGGCGTATCTTCTTATATGCTTGTGAGCT 800
267 uPheAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMetS 284
801 TTTTCGACTTGAACGGATATCTTCACTGTTGTTTCTGTGGTATTGTGATGT 850
284 erHisTyrThrTrpHisAsnValThrGluSerSerArgIleThrThrLys 300
851 CCCATTACACATGGCACAAATGTAACGGAGAGCTCAAGATTAACAACAAAG 900
301 HisThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTy 317
901 CATACCTTTTGAACCTTTGTCTCATTTCTTTCGGAGACATTTATTTCTTGTA 950
317 rValGlyMetAspAlaLeuAspIleAspLysTrpArgSerValSerAspT 334
951 TGTGGAAATGGATGCTTGGACATTGACAGTGGAGATCCGTGAGTGACA 1000

334 hrProGlyThrSerIleAlaValSerSerIleLeuMetGlyLeuValMet 350
1001 CACCGGGAACATCGATCGAGTCAGTCAATCTTAATGGTCTGTCATG 1050
351 valGlyArgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAl 367
1051 GTTGAAGACGACCGTTCGTTCCGTTATCGTTTCTATCTTAACCTAGC 1100
367 alyLysAsnGlnSerGluLysIleAsnPheAsnMetGlnValValIleT 384
1101 CAGAAGAATCAAGCGAGAAATCACTTTAATGTCAGGTGTGATTT 1150
384 rptPsrGlyLeuMetArgGlyAlaValSerMetAlaLeuAlaTyrAsn 400
1151 GGTGGTCTGCTCATGAGAGTGTGTTATCTATGCTCTTCATACAAC 1200
401 LysPheThrArgAlaGlyHisThrAspValArgGlyAsnAlaIleMetI 417
1201 AAGTTTACAAGGCGCGGCACACAGATGTACGCGGATGCAATCATGAT 1250
417 eThrSerThrIleThrValCysLeuPheSerThrValValPheGlyMetL 434
1251 CAGGATGACATAACTGCTGCTTTTATGACACAGTGGTGTGTTGATGC 1300
434 euThrLysProLeuIleSerTyrLeuLeuProHisGlnAsnAlaThrThr 450
1301 TGACCAACCACTCATAGCTACCTATTACCGCACACAGAACGCCACAG 1350
451 SerMetLeuSerAspAspAsnThrProLysSerIleHisIleProLeuLe 467
1351 AGCATGTTATCTGATGACAAACACCCCAAAATCCATACATATCCCTTGT 1400
467 uaspGlnAspSerPheIleGluProSerGlyAsnHisAsnValProArgp 484
1401 GGACCAAGACTCGTTCATTGAGCTTCAGGAACCAACAATGTCCTCGGC 1450
484 roAspSerIleArgGlyPheLeuThrArgProThrArgThrValHistyr 500
1451 CTGACAGTATACGTGGCTCTTTCGACACGCCCACTCGAACCGTCGATTAC 1500
501 TyrTrpArgGlnPheAspAspSerPheMetArgProValPheGlyGlyAr 517
1501 TACTGGAGACAATTTGATGACTCTTCATCGACCGCTCTTTGGAGGTG 1550
517 gGlyPheValProPheValProGlySerProThrGluArgAsnProProA 534
1551 TGCTTTGTACCTTTGTTCCAGTTCTCCAACTGAGAGAAACCCCTCTG 1600
534 spLeuSerLysAla 538
1601 ATCTTAGTAAGGCT 1614

seq_name: gb_p11:AF056190

seq_documentation_block:

LOCUS AF056190 1619 bp mRNA PLN 06-JUN-2000

DEFINITION Arabidopsis thaliana Na⁺/H⁺ exchanger (NHX1) mRNA, complete cds.

ACCESSION AF056190

VERSION AF056190.1 GI:5650176

KEYWORDS

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;

Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 1619)

AUTHORS Quintero,F.J., Blatt,M.R. and Pardo,J.M.

TITLE Functional conservation between yeast and plant endosomal

Na(+)/H(+) antiporters

JOURNAL FEBS Lett. 471 (2-3), 224-228 (2000)

MEDLINE 20231718

REFERENCE 2 (bases 1 to 1619)

1 MetLeuAspSerLeuValSerLysLeuProSerLeuSerThrSerAspHis 17
 3 ATCTTGGATTCTCTAGTGTGCGAAACTGCCCTTGCTGTATCGACATCTGATCA 52
 17 sAlaSerValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34
 53 CGCTTCTGTGGTTCGGTTGGAATCTCTTTGTGACACTTCTTTGTGCTTGGTA 102
 34 leValLeuGlyHisLeuLeuGluAsnArgTrpMetAsnGluSerIle 50
 103 TTGCTTCTTGCTCATCTTTTGGGAAGACAAATAGATGGATGAACCAATCCATC 152
 51 ThrAlaLeuLeuLeuLeuGlyLeuGlyThrGlyValThrIleLeuLeuIle 67
 153 ACCGCCCTTGTTGATTGGGCTGAGCACTGGTGTGTACCATTTTGTGTATTAG 202
 67 rLysGlyLysSerSerHisLeuLeuValPheSerGluAspLeuPhePheI 84
 203 TAAAGGAAAAAGCTCGCATCTTCTCGCTCTTTAGTGAAGATCTTTTCTTCA 252
 84 lefYrLeuLeuProProIleIlePheAsnAlaGlyPheGlnValLysLys 100
 253 TATATCTTTTGCACCACCATTATATTCAATGCAGGGTTTCAAGTAAAAAAG 302
 101 LysGlnPhePheArgAsnPheValThrIleMetLeuPheGlyValaValGI 117
 303 AAGCAGTTTTTCCGCAATTCGTGACTATTATGCTTTTTTGGTGTCTGG 352
 117 yThrIleIleSerCysThrIleIleSerLeuGlyValThrGlnPhePhe 134
 353 GACTATTATTCTTGGCAATCATATCTCTAGTGTGTAAACAGTCTCTTTA 402

134	ySlysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleGly	150
403	AGAAGTTGGACATTGGAAACCTTTGACTTTGGGTGATTATCTTGGCTATTTGGT	452
151	AlaIlePheAlaIaThrAspSerValCysThrLeuGlnValLeuAsnG1	167
453	GCCATATTTCCTGCAACAGATTACGTGTGACACATGCAAGGTTCGAATCA	502
167	nAspGluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValVala	184
503	AGACGAGACACCTTTGCTTTACAGCTCTTGATTCCGAGAGGGTGTGTGA	552
184	snAspAlaThrSerValValValPheAsnAlaIleGlnSerPheAspLeu	200
553	ATGATGCAACGTCAGTTGTGTGGTCTTCAACGCGATTTCAGAGCTTTGATCTC	602
201	ThrHisLeuAsnHisGluAlaIaPheHisLeuLeuGlyAsnPheLeuTy	217
603	ACTCACTTAACCCAGAAAGCTGCTTTTCATCTTCTTGAAACATCTTGTGA	652
217	rLeuPheLeuLeuSerThrLeuLeuGlyAlaAlaIaThrGlyLeuIleSerA	234
653	TTTGTTCCTCAAGTACCTTCCTGGTGTCTCAACCGCTCTGATAAGTG	702
234	laTyrValIleLysLysLeuTyrPheGlyArgHisSerThrAspArgGlu	250
703	CGTATGTTATCAAGAAGCTATACTTTGGAAAGCACTCAACTCAGCGAGAG	752
251	ValAlaLeuMetMetLeuMetAlaIaTyrLeuSerTyrMetLeuAlaGluLe	267
753	GTTCGCCCTTATGATGCTTAAGCGATCTTCTTATATGCTTGTGCTGAGCT	802
267	uPheAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValImets	284
803	TTTTCGACTTGACGGGTATCCTCAGCTGTTTTCCTGCTGCTATTTGTGATGT	852
284	exHisTyrThrTrpHisAsnValThrGluSerSerArgIleThrThrLys	300
853	CCATTACACATGGCAACATGTAAACGGAGAGCTCAAGAAATAACAACAAG	902
301	HisThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTy	317
903	CATACCTTTGCAACTTTGTTCATTTCTTCGGGAGACATTTATTTTCTTGA	952
317	rValGlyMetAspAlaLeuAspIleAspLysTyrArgSerValSerAspT	334
953	TGTTGGAATGGATGCCCTCGACATTCGACAAGTGGAGATCCGTGAGTGACA	1002
334	hrProGlyThrSerIleAlaValSerSerIleLeuMetGlyLeuValMet	350
1003	CACCGGGAACATCGATCGCAGTCAGCTCAATCCTAATGGTCTGGTCAATG	1052
351	valGlyArgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAl	367
053	GTGTGAAGACAGCGTTCGTCCTTCGTTATCGTTCTATCTCAACTTACG	1102
367	aLysLysAsnGlnSerGluLysIleAsnPheAsnMetGlnValIleLeu	384
1103	CAGAAGAAATCAAGCGACAAAATCAACTTTAACATGCGAGTTGTGATTT	1152
384	rpTrpSerGlyLeuMetArgGlyAlaValSerMetAlaLeuAlaTyrAsn	400
1153	GGTGGTCTGCTCATGAGAGGTGCTGTATCTATGGCTCTTGCAATACAAC	1202
401	LysPheThrArgAlaGlyHisThrAspValArgGlyAsnAlaIleMetI1	417
1203	AAGTTTACAAGGGCGGGACACAGATGTACCGCGGAATGCATCATCATGAT	1252
417	eThrSerThrIleThrValCysLeuPheSerThrValValPheGlyMetL	434
1253	CACGAGTACGATAACTGCTGCTCTTTTACACAGTGGGTGTGGTATGC	1302
434	euthrLysProLeuIleSerTyrLeuLeuProHisGlnAsnAlaThrThr	450

|||||TGACCAACCACTCATAAGCTACCTATTACCGCACCAACGCCACACG 1352
451 SerMetLeuSerAspAsnThrProLysSerIleHisIleProLeuLe 467
|||||AGCATGTTATCTGATGACACACACCCCAATCCATACATATCCCTTGT 1402
467 uAspGlnAspSerPheIleGluProSerGlyAsnHisAsnValProArgP 484
|||||GGACCAAGACTCGTTCATTGAGCTTCAGGGAACCAATGTGCTCGGC 1452
484 rAspSerIleArgGlyPheLeuThrArgProThrArgThrValHisTyr 500
|||||CTGACAGTATACGTGGCTCTTGACACGCGCCACTCGAACCGTGCAATTAC 1502
501 TyrTrpArgGlnPheAspSerPheMetArgProValPheGlyGlyAr 517
|||||TACTGGAGACAAATTTGATGACTCTTCATGCGACCCGCTCTTTGGAGTCG 1552
517 gGlyPheValProPheValProGlySerProThrGluArgAsnProProA 534
|||||TGGCTTTGTACCCCTTCTGTCAGGTTCTCCAACCTGAGAGAAACCCCTCTG 1602
534 spLeuSerLysAla 538
|||||ATCTTAGTAGAGCT 1616

seq_name: gb_p13:TM021B04

seq_documentation_block: 90019 bp DNA PLN 12-JUN-1997
LOCUS TM021B04
DEFINITION Arabidopsis thaliana BAC TM021B04.
ACCESSION AF007271
VERSION AF007271.1 GI:21911181
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS Arabidopsis thaliana
TITLE Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 90019)
Dante, M., Wamsley, P. and Gibson, A.
The sequence of A. thaliana TM021B04
Unpublished (1997)
REFERENCE
AUTHORS Washington University Genome Sequencing Center.
TITLE The A. thaliana Genome Sequencing Project
Unpublished (1997)
REFERENCE
AUTHORS Waterston, R.
TITLE Direct Submission
COMMENT Submitted (06-JUN-1997)

Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63108, USA
e-mail: rwlison@watson.wustl.edu

MAPPING: Clones were assigned to the YAC map by hybridization by
M. Lohdi, Cold Spring Harbor Laboratories, and fingerprinted
by M. Marra, WashU, to pick the best candidates for sequencing.

NOTICE: This sequence may not be the entire insert of this clone.
It may be shorter because we only sequence overlapping sections
once, or longer because we provide a small overlap between
neighboring submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate
chemistry; an attempt was made to resolve all sequencing problems,
such as compressions and repeats; all regions were covered by

sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

The actual start of this clone is at base position 1 of TM021B04 ;
the actual end is at base position 90019 of TM021B04. The
orientation of this clone is unknown.

NOTES:

Coding sequences below are predicted from computer analysis, using
the program Genefinder(P. Green and L. Hillier, ms in preparation).

FEATURES
source

Location/Qualifiers
1. 90019
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="V"
/clone="TM021B04"
/map="5"
1802. 3685
/gene="A.TM021B04.11"
join(1802. 2050,2468. 2686,2758. 2844,2901. 3095,
3275. 3685)
/gene="A.TM021B04.11"
/note="contains similarity to Synecococcus PCC7942
chromosomal region used as basis of neutral siteir
recombinational cloning vector (PID:g1174192)"
/codon_start=1
/evidence=not_experimental
/protein_id="AAB61082.1"
/db_xref="GI:2191197"
/translation="MVMLHSGVLI PCSNQKSF LPHSYRVCIVCSSETGLSIRRO
ALEQVDSKLSGSDERAA LSLVDLQKPDLCRCFGAARQVPPORLYLTLEELKNGINAA
SLSDPTDTTIGSTIERNLQAASGGIVAWKAEDLSQQLFTLTGFMFLWTLDLVSNF
GGTGLSVLDTTGTHTFSQYHNRVQKHYYII FHWTYCELRQHEAGHFLVAYLVGLPR
GYTLLSLALQKEGSLNIQAGSAFVDYEFLEFPNKKLCILLFONQMLNRFSCIALGVA
TEVLLYXAGAGGLDDISKYSFLLPLKNSDDYVNMLYGFVLMEOGLDGLVKSIGFTQKK
ADSOVRWSVLNTILLRRHEIARSKLAQAAMSGESVSGCIQIETSIDPSDI"
complement(3906. 4727)
/gene="A.TM021B04.14"
complement(join(3906. 4337,4418. 4497,4601. 4727))
/gene="A.TM021B04.14"
/note="coded for by A. thaliana cDNA T22079"
/codon_start=1
/protein_id="AAB61076.1"
/db_xref="GI:2191191"
/translation="MEATSLSSAATIISSSSPLSIFSPKRTDSSPPPRIVRLSNKK
EDKDYDPQHSNSSLFRNRTLSNDEAMGLVLSAASVKGWTTGSMGSPAPKTDIT
DIVSTFPWSLFTKSPRRMRVAFTCNVCGORTTALNPHAYTDGTVFVQCCGCVFHK
LVDNLNLFHEVAYVSSSFDYTDKAWDYSGLNLFDEDDDDNAGDSNDVFPPL"
complement(12650. 16625)
/gene="A.TM021B04.10"
complement(join(12650. 13129,13415. 13509,14187. 14498,
14596. 16014,16127. 16625))
/gene="A.TM021B04.10"
/note="similar to N. tabacum membrane-associated
salt-inducible protein (PID:g473874)"
/codon_start=1
/evidence=not_experimental
/protein_id="AAB61077.1"
/db_xref="GI:2191192"
/translation="MKSDFLTSTHNPISFIPKIPSRNSRISIKSSSSSSKVRPDPP
SLSDGNPEKPKRYEPKPHI LDDARRI LKKKAYLSLRLRNQSQAMT PKWIRTP
EONQVLEDDRGMQYGRVAAIKTVRGLSROGSDDMRFVSSFVAKUSFDMCVACG
VLKEQRCFCVKQLSVPVSVVYTVILRYGVQVKIKMAEETFLEMVECEPDVACG
TMCITYARWRHSAMLTFFYKQERRILLSTSVNFMSSLQKSGHKGKVIDLWEMV
EEGVPPNEFTYLVSSYAKQGFKEALKAFGEMKSLGFVPEVTVSSVLSLVKAGD
WKAIGLYEDMRSGQVPSNYTCATMISLYYKTEVNPALSLFADMERNK IPADDEVIR
GLIIRYGLGLGFHDQASMEETERNLLADEKTYLAMSQVHLNSGVMKALDIETMM
KTRDIPLSRFATVIMLQCYAKIQNVDCAEAFRALSKGLPDPASSCNMDNLTYRLNL


```
95 .....ThrGlyLeuIleSerAlaTyrValIleLysLysLeuTyrPheG1 243
66474 TTGCAGACCGGTCGATAAGTGCATGTTATCAAGAAAGCTATACTTTGG 66425
243 Y..... 243
66424 AAGGTACATATTTAAAGTCATATATGTGACTTGTCTTTTAAAGTTGTTT 66375
244 .....ArgHisSer 246
66374 CCTGTTGGTCTAAGGTAGATGGCTCACATTAATGTGCACAGGCACATCA 66325
247 ThrAspArgGluValAlaLeuMetMetLeuMetAlaTyrLeuSerTyrMe 263
66324 ACTGACCGAGAGGTTGCCCTTATGATGCTTATGGCGTATCTTCTTATAT 66275
263 tleuAlaGluLeu..... 267
66274 GCTTGTGAGGTTAGATTCTTTTAAATAACACGAAACATATTCATTCAGTT 66225
268 .....Phe 268
66224 GCATTCAACACACAAATTATATGTGATTATTAATTTGGTTGCAGCTTT 66175
269 ..AspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMetSer 284
66174 TCGACTTGAGCGGTATCTCCTACTGTGTTTTTCTGTGGTATTGTGATGCC 66125
285 HisTyrThrTrpHisAsnValThrGluSerSerArgIleThrThr..... 299
66124 CATTACATGCGCACAAATGTACGAGGAGCTCAAGAATAACAACAAGTA 66075
299 ..... 299
66074 TCCAAACTTTAGTTTCTCGGCTTTTGTGAGATTAAACACTTTTACCTC 66025
300 .....Lys..HisThrPheAlaT 305
66024 CCAATGATTATCTTATGATTTTTTCGATGAACAGGCATACCTTTGCAA 65975
305 hrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrValGlyMetAsp 321
65974 CTTTGTCAATCTTCGGGAGACATTTATTTCTGTATGTTGGAATGGAT 65925
322 AlaLeuAspIleAspLysTrpArgSerValSerAspThrProGlyThrSe 338
65924 GCCTTGGACATTGACAAGTGGAGATCCGTGAGTGACACACCGGGAACATC 65875
338 rIleAlaValSerIleLeuMetGlyLeuValMetValGlyArgAlaA 355
65874 GATCGAGTGAGCTCAATCCTAATGGGTCTGGTCATGGTTGGAAGAGCAG 65825
355 laPheValPheProLeuSerPheLeuSerAsnLeuAlaLysLysAsnGln 371
65824 CGTTGCTCTTTCCGTTATCGTTTCTATCTAATCTAGCCAAAGAAGTCAA 65775
372 SerGluLysIleAsnPheAsnMetGlnVal.Val..... 382
65774 ACGGAGAAATCAACTTTAACATGCGAGGTATGATCATCTACTTCTGGTAC 65725
382 ..... 382
65724 TAATAGATTCAATAATCTCTCGGATTTTCAGCTATTTATATCTATCAATTT 65675
383 .....IleTrpTrpSerGlyLeuMetArgGlyAl 392
65674 TTTCCTCAATCTTTAGGTTGTGATTTGGTGGTCTGTCTCATGAGAGGTGC 65625
392 aValSerMetAlaLeuAlaTyrAsnLys..... 401
65624 TGTATCTATGGCTCTTGCATACACAAGGTATAAAAGAAATGCTCTTAA 65575
401 ..... 401
```

```

65574 TACTCTCTATGAAGCTTCAAAATGACAAATCATCATTTAAACACTGTGTT 65525
402 .....PheThrArgAlaGlyHisThrAspValArgGlyAsnAlaIle 415
65524 CTATTGTAGTTTACAAGGCGCGGCACACAGATGTACGCGGAATGCAAT 65475
415 eMetIleThrSerThrIleThrValCysLeuPheSerThrVal..... 429
65474 CATGATCAGAGTACCATCACTGTCTCTCTTTTACGACAGTGGTAAGAA 65425
429 ..... 429
65424 ATCTCTGGATGATCAAAATGAGTTACATAATATTTCTTAATGAATGGAC 65375
430 .....ValPheGlyMetLe 434
65374 ATTTGGTTAATAGATTCCTTGAACCTTCTTTTCAGGTTGTTGGTATGCT 65325
434 uThrLysProLeuIleSerThrLeuLeuProHisGlnAsnAlaThrThrS 451
65324 GACCAACACCATCATAAAGTACTATTACCGCACCAAGAACGCCACACGA 65275
451 eMetLeuSerAspAspAsnThrProLysSerIleHisIleProLeuLeu 467
65274 GCATGTATCTGATGACACACCCCAAAATCCATACATATCCCTTTGTTG 65225
468 AspGlnAspSerPheIleGluProSerGlyAsnHisAsnValProArgPr 484
65224 GACCAAGACTCGTTTCATTGAGCCTTCAGGGAACACCAATGTGCCGCGCC 65175
484 oAspSerIleArgGlyPheIleuThrArgProThrArgThrValHisTyrT 501
65174 TGACAGTATACGTCGCTTCTTGACAGCGCCACTCGAACCGTGCATTTACT 65125
501 yTrpArgGlnPheAspAspSerPheMetArgProValPheGlyGlyArg 517
65124 ACTGGAGACAATTTGATGACTCCTTCATCGACCGCCGCTTTTGGAGTCTG 65075
518 GlyPheValProPheValProGlySerProThrGluArgAsnProProAs 534
65074 GGCTTTGTACCTTTGTTCCAGGTTCTCCAACTGAGAGAAACCCCTCTGA 65025
534 pLeuSerLysAla 538
65024 TCTTAGTAAGGCT 65012
seq_name: gb_p11:AB038492
seq_documentation_block:
LOCUS AB038492 2637 bp mRNA PLN 18-AUG-2000
DEFINITION Atriplex gmelini AgNHX1 mRNA for Na/H antiporter Nhx1, complete cds.
ACCESSION AB038492
VERSION AB038492.1 GI:9857313
KEYWORDS Na/H antiporter Nhx1.
SOURCE Atriplex gmelini cDNA to mRNA.
ORGANISM Atriplex gmelini
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae; Caryophyllales; Chenopodiaceae; Atriplex.
REFERENCE
AUTHORS Hamada,A., Hayakawa,T., Shono,M., Xia,T., Hayashi,Y. and Tanaka,A.
TITLE Isolation and characterization of a Na/H antiporter gene from the halophyte Atriplex gmelini
JOURNAL Unpublished (2000)
REFERENCE
AUTHORS Hamada,A., Hayakawa,T., Hamada,A. and Tanaka,A.
TITLE Direct Submission
JOURNAL Submitted (17-FEB-2000) to the DDBJ/EMBL/GenBank databases.
Takahiko Hayakawa, Plantech Research Institute, Research Center; 1000 Kamoshida-cho, Aoba-ku, Yokohama, Kanagawa 227-0033, Japan
(E-mail:pr10012@cc.m.kagaku.co.jp, Tel:+81-45-963-3520,

```

```

FEATURES
source
Location/Qualifiers
1..2637
/organism="Atriplex gmelini"
/db_xref="taxon:118074"
682..2349
/gene="AgNHX1"
682..2349
/gene="AgNHX1"
/codon_start=1
/product="Na/H antiporter Nhx1"
/protein_id="BAB11940.1"
/db_xref="GI:9857314"
/translation="MMSQLSSLSGKMDALTTSDHASVSVSMNLFVALLGCGCIVIGHLL
ENRMNESITALLIGLATGVVILLISGGKSHLLVFSDELFFIYLLPPIIFNAGFQV
KKQFFNFITIVLFAGVGLVFTIISLGALSIFKKLDIGTLELADYLAIGAIFAAT
DSVCTLOVNODETPLLYSLVFGVGVNDATSVLFNATQSPDLTRIDHRIALQPMGN
FLXLFIASTILGAFTPLLSAYIILKLYFGRHSTDREVALMMLMAYLSYMLAEFLYLSG
ILTVFCGIVMSHYTHWNTESRVTTKHAFATLSFAVFLFLVYGMALDIEKWRFL
VSDPGISVAVSSILGLVWVGRAAFVFLPUSLWLNFAKKSQSEKVTFFNOQIVLWAGL
MRGAVSMALAYNOPTRSRGTOLRGNAIMITISTISVLFSTWVGLLTKPLIMFLDPOP
KHFTSCSTVSDGSPKSYSLPLLEGNDYEDVGNHEDTTPRTIVRPSSLRMLLN
APTHTVHYWRKFFDDSEMRPVFGGFGFVPGSPTEQSTNNLVDR"
2607
polyA_site
/note="30 a nucleotides"
BASE COUNT 685 a 509 c 530 g 913 t
ORIGIN
alignment_scores
Quality: 2168.50 Length: 547
Ratio: 4.286 Gaps: 5
Percent Similarity: 92.505 Percent Identity: 77.697
alignment_block
US-09-271-584A-2 x AB038492 ..
Align seg 1/1 to: AB038492 from: 1 to: 2637
2 LeuAspSerLeuValSer...LysLeuProSerLeuSerThrSerAspHis 17
694 TTAAGCTCTTTGTTGTTCTGGAAAGATGGACGCTCTTACCACCTTCTGATCA 743
17 sAlaSerValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34
744 CGCTTCTGTGCTCGATGAACCTGTTTCTGGCACGTGTTATGTTGTTGTA 793
34 leValLeuGlyHisLeuLeuGluAsnArgTrpMetAsnGluSerIle 50
794 TCGTAATTGGTCTATCTTAGAGGAGAAATCGTTGGATGAATGAGTCCATC 843
51 ThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeuLeuIleSe 67
844 ACTGCCCTCTTATAGTTTGGCTACTGGGTGTGATTCCTGCTGATTAG 893
67 rLysGlyLysSerSerHisLeuValPheSerGluAspLeuPhePheI 84
894 TGGAGGAAAAAGTTCACATCTTTGGTCTTCAGTGAAGATCTTTCTTCA 943
84 leThrLeuLeuProIleIlePheAsnAlaGlyPheGlnValLysLys 100
944 TATACCTTCTTCCACCATATATTCATGCAGGCTTCAGGTGAGGAAG 993
101 LysGlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaValGl 117
994 AAGCAGTCTTCCCAACTTCATTACTATTGTTGAGCTGTTGG 1043
117 yThrIleIleSerCysThrIleIleSerLeuGlyValThrGlnPhePheL 134
1044 TACATTGGTATCATTCACCATCATATCTCTGGGAGCGGTGTCAATTTT 1093
134 ysLysLeuAspIleGlyThrPheAspLeuGlyAspThrLeuAlaIleGly 150

```

```
1094 AGAAATTGGATATTGGTACTCTGGAGTTGGCAGACTATCTTGCAATTGGT 1143
151 AlalPheAlaLathrAspSerValCysThrLeuGlnValLeuAsnG1 167
|||||
1144 GCAATATTCTGCTGCACAGATTCTGTTGTCACACTGCAGGTTCTTAATCA 1193
167 nAspGluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValVala 184
GGATTGAGACCCCTCTGCTACAGCTTGGTCTTTGGCGAGGGTGTGTATA 1243
184 snAspAlaThrSerValValValPheAsnAlaLleGlnSerPheAspLeu 200
|||||
1244 ATGATGCCACATCAGTGGTCTCTTTTCAATGCAATTCAGACCTTGACCTC 1293
201 ThrHisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTy 217
|||||
1294 ACAAGAATTGATCACAAGATAGCTTTACAATTTATGGCAACTTCTTATA 1343
217 rLeuPheLeuLeuSerThrLeuLeuGlyAlaAlaThrGlyLeuLeuSerA 234
|||||
1344 TTTATTATTCGAAGCAGCATCTTGGAGCATTTACTGGCTTGCTCAGTG 1393
234 laTyrValLleLysLeuTyrPheGlyArgHisSerThrAspArgGlu 250
|||||
1394 CTTACATTATCAAAAGCTGTACTTGGAGGCATTCACACTGATCGTGAG 1443
251 ValAlaLeuMetMetLeuAlaTyrLeuSerTyrMetLeuAlaGluLe 267
|||||
1444 GTTGCTTTAATGATGCTATGCGCTTATCTATCTTACATGCTGCTGAAC 1493
267 upAspLeuSerGlyLleLeuThrValPhePheCysGlyLleValMets 284
|||||
1494 TTTCTATTGAGTGAATCTTACTGCTATCTTCTGTGGGATTCATGT 1543
284 erHisTyrThrTrpHisAsnValThrGluSerSerArgLleThrLys 300
|||||
1544 CCCATTATACCTGGCACAATGTGACAGAGACTCAAGAGTAAACCAACAG 1593
301 HisThrPheAlaThrLeuSerPheLeuAlaGluThrPheLlePheLeuTy 317
|||||
1594 CATGCTTTTGAACACTATCTTTTGTGCTGAGGTTTCTCTTATTCTATA 1643
317 rValGlyMetAspAlaLeuAspLleAspLysTrpArgSerValSerAspT 334
|||||
1644 TGTCGGTATGATGATGGACATTTGAGAGTGGAGATTTGTTAGTGATA 1693
334 hrProGlyThrSerLleAlaValSerSerLleLeuMetGlyLeuValMet 350
|||||
1694 GTCCTGGCATTTCTGCTGAGTTCATATTTGCTAGGCTAGTCATG 1743
351 ValGlyArgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAl 367
|||||
1744 GTTGAAGAGCAGCTTTGTTTTCCTCTATCTGTTGATGAACCTTTCG 1793
367 alyLysAsnGlnSerGluLysLleAspPheAsnMetGlnValValLleT 384
|||||
1794 CAAGAAATCGCAAGTGAAGAGTCACTTTCAACCGAGATTTGTCATAT 1843
384 rpTrpSerGlyLeuMetArgGlyAlaValSerMetAlaLeuAlaTyrAsn 400
|||||
1844 GTTGGGCTGGTCTTATGAGAGTGTGTTTCCATGCGCACTTGTATATAT 1893
401 LysPheThrArgAlaGlyHisThrAspValArgGlyAsnAlaLleMetLl 417
|||||
1894 CAGTTTACGAGGCTGCGGCACACACAGCTTAGGGGAAATGCAATCATGAT 1943
417 eThrSerThrLleThrValCysLeuPheSerThrValValPheGlyMetL 434
|||||
1944 CACAAGCACTATATCTGTTGCTTTTAGTACAATGGTGTGCGGGTGC 1993
434 euThrLysProLeuLleSerTyrLeuLeuProHisGlnAsnAlaThrThr 450
|||||
1994 TGCAAGAGCTCTCATCATGTTTGTCTGCTCAACCGAACACACTTCATC 2043
```

```
451 .....SerMetLeuSerAspAspAsnThrProLysSerLleHisIlePr 465
|||||
2044 AGTTCACACACCGTATCAGATGTGGAGAGTCCAAAGTCATACTCGTTGCC 2093
465 oleuLeuAsp.....GlnAspSerPheIleGlu....ProSerGlyAsnH 479
|||||
2094 ACTCCTTCAGGGCAACCAAGATTATGAAGTTGATGTGGGAAACGAAACC 2143
479 is.....AsnValProArgProAspSerIleArg 488
|||||
2144 ATGAAGACACCACTGAGCGCGGACTAGTTCGACCTAGTAGCCTCGGC 2193
489 GlyPheLeuThrArgProThrArgThrValHisTyrTyrTrpArgGlnPh 505
|||||
2194 ATGCTTCATAAATGCACCTACTCACACCGTCCATCACTATTGGCGCAANT 2243
505 eAspAspSerPheMetArgProValPheGlyGlyArgGlyPheValProp 522
|||||
2244 CGATGACTCTTCATGCGCGCGTGTGTTGGTGGCGGGTGTGTACCTT 2293
522 heValProGlySerProThrGluArgAsnProProAspLeu 535
|||||
2294 TTGTCCTCGGGTTACCTACTGAACAAAGCACCACCAATTTG 2334
seq_name: gb_p11:AB033989
seq_documentation_block: 2218 bp mRNA PLN 06-OCT-2000
LOCUS AB033989
DEFINITION Ipomoea nil Purple mRNA for Na+/H+ exchanger, complete cds.
ACCESSION AB033989
VERSION AB033989.1 GI:10716128
KEYWORDS
SOURCE
ipomoea nil (strain:KZSK-2) fully colored flower buds cdna to
mRNA, clone_lib:cdna library.
ORGANISM
Ipomoea nil
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
I; Solanales; Convolvulaceae; Ipomoea.
REFERENCE
1 (sites)
AUTHORS Fukada-Tanaka,S., Inagaki,Y., Yamaguchi,T., Saito,N. and Iida,S.
TITLE Colour-enhancing protein in blue petals
JOURNAL Nature 407, 581 (2000)
REFERENCE 2 (bases 1 to 2218)
AUTHORS Fukada-Tanaka,S., Inagaki,Y., Yamaguchi,T. and Iida,S.
TITLE Direct Submission
JOURNAL Submitted (26-OCT-1999) to the DDBJ/EMBL/GenBank databases.
Yoshishige Inagaki, National Institute for Basic Biology, Division
of Gene Expression and Regulation 1; Myodaiji-cho, Okazaki, Aichi
444-8585, Japan (E-mail:yinagaki@nibb.ac.jp, Tel:81-564-55-7684,
Fax:81-564-55-7685)
FEATURES
Location/Qualifiers
source
1..2218
/organism="Ipomoea nil"
/strain="KZSK-2"
/db_xref="taxon:35883"
/clone_lib="cdna library"
/tissue_type="fully colored flower buds"
300..1928
/gene="Purple"
300..1928
/gene="Purple"
/codon_start=1
/product="Na+/H+ exchanger"
/protein_id="BAB16380.1"
/db_xref="GI:10716129"
gene
CDS
300..1928
/gene="Purple"
300..1928
/gene="Purple"
/codon_start=1
/product="Na+/H+ exchanger"
/protein_id="BAB16380.1"
/db_xref="GI:10716129"
/translation="MAFGSLQLQNSDLFTSDHASVSVSMNLFVALLCACIVLGHLEE
NRWNESITALTIGCTGWTLLSGKSHLLVSEDLFFIYLLPPIFNAGFOVKK
KOFFVNEWTIMLFGAIGTLISCSISFGAVKIFKHLOIDFDFDGYLAIGAIFATDS
VCTLOVLSODETPLLYSLVFGGVNDATSVLENAIQSFDMTSFDPKIGLHFTGNFL
YLFLSTFLGVGIGILLCAIILKKLYFGRHSTDRVALMMLMSYLSYIMAEFLYLSGIL
TVFFCGIVMSHYTHNVTESRKVTRHRSFATLSLNIAKNSDKISFRQOIIIIWAGLMR
NSQGLSVAVSSILVGLILVGRAAFVFPFLSLNIAKNSDKISFRQOIIIIWAGLMR
```

GAVALAYNKFTTSGHTSLHENAIMITSTVTVVFLSTVVFGLMTPKPLINLLPPPHKQ
MPSGHSNMTSESPSKHFTVPLIDNPDSESDMITGPEVARPTALRMILLRTPTHTVH
RYWRKFDSEMRPVPFGRGFPVAGSPVQSPR"
BASE COUNT 568 a 439 c 492 g 719 t
ORIGIN

alignment_scores:
Quality: 2041.00 Length: 541
Ratio: 4.182 Gaps: 4
Percent Similarity: 90.203 Percent Identity: 73.752
alignment_block:
US-09-271-584A-2 x AB033989 ..
Align seg 1/1 to: AB033989 from: 1 to: 2218
2 LeuAspSerLeuValSerLysLeuProSerLeuSerThrSerAspHisAl 18
312 TTGCTCTCTTTGCTGCAAAAT...TCGGATTGTTACGCTCGATCATGC 358
18 aSerValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysIleVal 35
359 TTCGGTTGTCGATGAACCTCTTTGCGCGFTGCTTTGCGCATGCATG 408
35 allLeuGlyHisLeuLeuGluGluAsnArgTIPMetAsnGluSerIleThr 51
409 TWTCTGGCCATCTACTCGAGGAGATCGCTGGGTGAACGAATCCATTACT 458
52 AlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeuLeuIleSerLys 68
459 GCCCTTATAATTTGTTGTCACCGGAGTTGTAATTTTGCCTCTAGCGG 508
68 sGlyLysSerSerHisLeuLeuValPheSerGluAspLeuPheIleThr 85
509 AGGAAGAGTTCCATCTCTCTGCTTTAGCGAAGATCTTTCTTTATAT 558
85 yrLeuLeuProIleIlePheAsnAlaGlyPheGlnValLysLysLys 101
559 ATCTCTGCCACCTATATATATTCATCGGGGTTTCAAGTGAANAAGAG 608
102 GlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaValIleGlyTh 118
609 CAGTTTTCGTGAACCTTCATACAATATATGCTGTTTGGAGCTATTCGCAC 658
118 rIleIleSerCysThrIleIleSerLeuGlyValThrGlnPhePheLysL 135
659 ACTTATTAGCTGTTCTATTATATCATTTGTTGCGGTCAAAATTTTCAAGC 708
135 ysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleGlyAla 151
709 ACTTAGACATTCACCTCTCGGATTTGGAGATATTTAGCAATTTGGTCGG 758
152 IlePheAlaAlaThrAspSerValCysThrLeuGlnValLeuAsnGlnAs 168
759 ATATTGCTGCAACCGGATTCCTGTTGACATTCAGGTGCTCAGTCAGGA 808
168 pGluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValValAsnA 185
809 TGAGACGGCCCTTACTTACGFTCTGCTGTTTGGAGAAGGGTTCGTCATG 858
185 spAlaThrSerValValValPheAsnAlaIleGlnSerPheAspLeuThr 201
859 ATGCTACATCTGCTGCTCTTTTATGCTATTCATCAAGTTTGGACATGACT 908
202 HisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTyrLe 218
909 AGTTTGTATCCAAAAATTTGGGCTTCATTTTCATTTGGAAACTTCTTGTA 958
218 uPheLeuLeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIleSerAla 235
959 ATTTCTCGAGCAGCTTTTGGGCGTGGGAATTGAGCTGCTTTGTGCTT 1008

235 yrValIleLysLysLeuTyrPheGlyArgHisSerThrAspArgGluVal 251
1009 ATATTATCAAAAGCTATACCTTTGGCAGGCACTCAACGATCGTGAGTT 1058
252 AlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPh 268
1059 GCCCTTATGATGCTCATGCTTACTTGTCTTATATATAATGGCCGAGTTATT 1108
268 eAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMetSerH 285
1109 CTATCTAAGCGGCATACCTTACTGTATTTCTGCTGGAATTTGTCATGCTC 1158
285 isTyrThrTrpHisAsnValThrGluSerSerArgIleThrThrLysHis 301
1159 ATTATACCTGGCACAATGTTACCGAGAGCTCAAGGTCACCTACTAGGCAT 1208
302 ThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrVa 318
1209 TCCTTTGCAACTCTGTCATTTGTCGAGAGACATTTATCTTCCTCTATGT 1258
318 lGlyMetAspAlaLeuAspIleAspLysTyrTrpArgSerValSerAspThrP 335
1259 TGGTATGATGCTTGGATATCGAAAAATGAAATTTGTGAAAAATAGTC 1308
335 roGlyThrSerIleAlaValSerSerIleLeuMetGlyLeuValMetVal 351
1309 AGGACCTATCAGTTGTCAGTGCATCAATATTTGGTAGGCTTAATCTTAGTA 1358
352 GlyArgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAlaLy 368
1359 GGCAGAGCTGGTTCGTATCCCTCTGCTTTTATCCAACTTAGCAAA 1408
368 sLysAsnGlnSerGluLysIleAsnPheAsnMetGlnValValIleTrpT 385
1409 GAAAACTCTTCGACAGATATCCTTTAGGCAACAAATAATAATTTGGT 1458
385 rpSerGlyLeuMetArgGlyAlaValSerMetAlaLeuAlaTyrAsnLys 401
1459 GGGTGGCTTAATGAGAGCGCCGCTCAATAGCACCTTGGCTATAATAAG 1508
402 PheThrArgAlaGlyHisThrAspValArgGlyAsnAlaIleMetIleTh 418
1509 TTTACAACCTCGGGGCATCGCTATTCACGAGAACCAATATGATTAC 1558
418 rSerThrIleThrValCysLeuPheSerThrValValPheGlyMetLeuT 435
1559 AAGTACTGTTACGGTTGTTCTGTTACACAGATTTGTTATTCGGGTTGATGA 1608
435 hrLysProLeuIleSerTyrLeuLeuPro..... 444
1609 CGAAGCCTCTGATAAACCCTTCTGCTACCCCGCACAGAGATGCCAAGC 1658
445 ...HisGlnAsnAlaThrThrSerMetLeuSerAspAsnThrProLy 460
1659 GGTCAATTCGTCAATGACAACA.....TCCGACCCAGTAGTCCGAA 1699
460 sSerIleHisIleProLeuLeuAspGln.....AspSerPheIleGluP 475
1700 GCATCTCAGGTCGCACTCTCTGGACAACCAACCTGACTCAGAAAGCGATA 1749
475 roSerGlyAsnHisAsnValProArgProAspSerIleArgGlyPheLeu 491
1750 TGATAACCGGACCTGAGGTTGCTGACCAACTGCTTTCGCGCATGCTGCTA 1799
492 ThrArgProThrArgThrValHisTyrTrpArgGlnPheAspAspSe 508
1800 AGGACGCCAACCCACACCGTCACCGCTACTGGCGTAAGTTTGTATGATTC 1849
508 rPheMetArgProValPheGlyArgGlyPheValProPheValProG 525
1850 GTTTATGCTGCCGCTGTTTGGCGGGGATTCGTTCCGTTTTCGCGG 1899
525 lySerProThrGluArgAsnPro 532

|||||.....
1900 GCTCACCAGTTGACGAGCCCT 1922

seq_name: gb_p11:AB021878

seq_documentation_block:
LOCUS AB021878 2313 bp mRNA PLN 12-AUG-1999
DEFINITION Oryza sativa OsNHX1 mRNA, complete cds.
ACCESSION AB021878
VERSION AB021878.1 GI:5731736
KEYWORDS
SOURCE Oryza sativa (cultivar:Nipponbare) cDNA to mRNA.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
Poaceae; Oryza.
REFERENCE
1 (sites)
Fukuda,A., Nakamura,A. and Tanaka,Y.
Molecular cloning and expression of the Na⁺/H⁺ exchanger gene in
oryza sativa
Biochim. Biophys. Acta 1446 (1-2), 149-155 (1999)
2 (bases 1 to 2313)
Fukuda,A., Tanaka,Y. and Nakamura,A.
Direct Submission
Submitted (25-DEC-1998) to the DDBJ/EMBL/GenBank databases.
Atsunori Fukuda, National Institute of Agrobiological Resources,
Department of Plant Physiology; Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan (E-mail:fukuda@abr.affrc.go.jp, tel:81-298-38-8376,
Fax:81-298-38-8347)
Location/Qualifiers
1. .2313
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
297..1904
/gene="OsNHX1"
297..1904
/gene="OsNHX1"
/note="similar to yeast sodium/proton exchanger"
/codon_start=1
/protein_id="BAA83337.1"
/db_xref="GI:5731737"
translation="MGMEVAAARLGAlyTTSYASVSYINLFVALLCACIVLGHLEE
KRWNEFTALIIGLCTGVWLLMTKGKSSHLFVSEDLFIYLLPLIENAGFOVKK
NQFENFTILFAGVCTMISFTTISAATAIFSRMNLGIDVGDFLAIGAFSATDS
VCTLOVMDTEPLYSILVFGGVNDATSVLNFALQNFDLVHIDRAVVKLFIGNFF
YLFUSSTFLGVFAGLLSAYIIKKLYIGRHSTVDREVALMMLMAYLSYMLAELLDLSGIL
TVFFCGIYMSHYTWHNVTESSRVTTKHAFATLSIAETFLFYVGMALDIEKWEFAS
DRPKSIGISILLGLIGRAAFVFLSFLSNLTKKAPNEKITRQOVVIMWAGLMR
GAVSTALAYNKTFRSGHTLHGNAIMITSTITVVFSTVMVGMKPLIKLLLPASCH
PVTSEPPSPKSLHSPLLTSMOGLSTNIVRPSLRMLLTKPTHVHYWRKFDDA
LMRPFGRGVFPSPSPSQSGGR"
2313
polyA_site
/note="17 a nucleotides"
BASE COUNT 543 a 538 c 569 g 663 t
ORIGIN

alignment_scores:
Quality: 2023.50 Length: 531
Ratio: 4.121 Gaps: 2
Percent Similarity: 92.467 Percent Identity: 73.258

alignment_block:

US-09-271-584A-2 x AB021878 ..

Align seg 1/1 to: AB021878 from: 1 to: 2313

2 LeuAspSerLeuValSerLysLeuProSerLeu...SerThrSerAspHi 17

.....:|||||.....

303 ATGGAGGTGGCGCGCGCTCTGTACACGACCTCCGACTA 352

17 sAlaserValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34
:|||||.....
353 CGCGTCGGTGTGTCCATCAACCTGTTCTGTCGCGCTGCTCTGCGCCTGCA 402
34 leValLeuGlyHisLeuLeuGluAsnArgTrpMetAsnGluSerIle 50
:|||||.....
403 TCGTCCTCGGCACCTCTCTCGAGGAGAAATCGCTGGTCAATGAGTCCATC 452
51 ThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeuLeuIleSe 67
:|||||.....
453 ACCGGCTCATATCGGGCTCTGCACCGCGGTGGTATCTTGTGCTGATGAC 502
67 rLysGlyLysSerSerHisLeuLeuValPheSerGluAspLeuPhePheI 84
:|||||.....
503 CAAAGGGAAGAGTCGCATATTATTCGCTTCAGTGAGGATCTCTTCTTCA 552
84 leTyrLeuLeuProProIleIlePheAsnAlaGlyPheGlnValLysLys 100
:|||||.....
553 TCTACCTCTCCCTCCGATCATCTTCAATGCAGGTTTTCAGGTAAGAAA 602
101 LysGlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaValGI 117
:|||||.....
603 AAGCAATTCTTCCGGAATTCATGACGATCACATATTATTTGAGCGCGTCGG 652
117 yThrIleIleSerCysThrIleIleSerLeuGlyValThrGlnPhePheL 134
:|||||.....
653 GACAATGATATCTTTTTCACAATATCTATTGCTGCCATTCGAATATTCA 702
134 ysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleGly 150
:|||||.....
703 GCAGAATGACATTTGGAACGCTGGATGTAGGAGATTTTCTTGCATTTGGA 752
151 AlaIlePheAlaAlaThrAspSerValCysThrLeuGlnValLeuAsnGI 167
:|||||.....
753 GCCATCTTTCTCGACACAGATCTGTCTGCACATTCAGGTCCTCAATCA 802
167 nAspGluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValValA 184
:|||||.....
803 GGATGAGACACCTTTTGTACAGTCTGGTATTCGGTGAAGGTGTGTGA 852
184 snAspAlaThrSerValValPheAsnAlaIleGlnSerPheAspLeu 200
:|||||.....
853 ACGATGCTACATCAATCTGCTTTTCAACGACATACAGAACTTTGATCTT 902
201 ThrHisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTy 217
:|||||.....
903 GTCCACATAGATCGCGCTGCTGTTCTGAAATCTCTGGGAACCTCTTTTA 952
217 rLeuPheLeuLeuSerThrLeuLeuGlyAlaAlaThrGlyLeuLeuSera 234
:|||||.....
953 TTTATTTTGTGAGCACCTTCTCTGGAGTATTTGCTGGATTGCTCAGTG 1002
234 laTyrValIleIleLysLysLeuTyrPheGlyArgHisSerThrAspArgGlu 250
:|||||.....
1003 CATACATAATCAAGAGCTATACATTGGAAGGCATCTTACTGACCGGTGAG 1052
251 ValAlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLe 267
:|||||.....
1053 GTTGCCCTTATGATGCTCATGGCTTACCTTTTCATATATGCTGGCTGAGTT 1102
267 uPheAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMetS 284
:|||||.....
1103 GCTAGATTTGAGCGGCATCTCACCGTATTTCTGTGTTGTTGTTGTTGTT 1152
284 erHisTyrThrTrpHisAsnValThrGluSerSerArgIleThrThrLys 300
:|||||.....
1153 CACATTACATTTGGCATAACGTCACAGAGAGTTCAAGAGATTACACAAAG 1202
301 HisThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTy 317
:|||||.....
1203 CACGATTTGCAACTCTGCTCTCATTCCTGAGACTTTTCTCTTCTTCTGTA 1252
317 rValGlyMetAspAlaLeuAspIleAspLysTrpArgSerValSerAspTr 334


```
|||||
1253 TGTGGGATGCATTCGATTTGATATTCGAAATGGGAGTTGGCAGTGACA 1302
334 hrProGlyThrSerIleAlaValSerSerIleLeuMetGlyLeuValMet 350
1303 GACCTGGCAATCCATTTGGGATGAAGTCAATTTGCTAGGATGGTCTG 1352
351 ValGlyArgAlaAlaPheValPheProLeuSerPheLeuSerAsnLeuAl 367
1353 ATTTGGAGAGCTGCTTTTGTATTCGCGCTGCTGTTCTTGTGCAACCTAAC 1402
367 aLysAsnGlnSerGluLysIleAsnPheAsnMetGlnValValIleT 384
1403 AAAGAAGGACCCGATCAATAAATAACCTGGAGACAGCAAGTTGTAATAT 1452
384 rPrpSerGlyLeuMetArgGlyAlaValSerMetAlaLeuAlaTyrAsn 400
1453 GGTGGGCTGGGCTGATGAGAGGAGCTGTCTCGATTCTCTCTTACAAAT 1502
401 LysPheThrArgAlaGlyHisThrAspValArgGlyAsnAlaIleMetI 417
1503 AAGTTTACAAGATCTGGCCATCTACTACGTCGACGGCAATGCAATATGAT 1552
417 eThrSerThrIleThrValCysLeuPheSerThrValValPheGlyMetL 434
1553 CACCAGCACCATCACCTGCTGCTCTTTTACACTATGTTATTTGGGATCA 1602
434 euThrLysProLeuIleSerTyrLeuLeuProHisGlnAsnAlaThrThr 450
1603 TGACAAAGGCATTGATGAGGCTGCTGCTACCGGCTCAGGCCATCTCTGC 1652
451 SerMetLeuSerAspAsnThrProLysSerIleHisIleProLeuLe 467
1653 ACC.....TCTGAGCTTCATCACCAAGTCCCTGCATCTCTCTCTCT 1696
467 uAspGlnAspSerPheIleGluProSerGlyAsnHisAsnValProArgP 484
1697 GACAAGCATGCAAGGTTCTGACCTCGAGAGTAGTACAACCAACATTTGGAG 1746
484 roAspSerIleArgGlyPheLeuThrArgProThrArgThrValHisTyr 500
1747 CTTCAGCTCGCGAGTCTCTACCAAGCCAGCCACACTCTCCACTAC 1796
501 TyrTrpArgGlnPheAspAspSerPheMetArgProValPheGlyGlyAr 517
1797 TACTGGCGCAAGTTCACACCGGCTGATCGCACCGATGTTTGGCGGGCG 1846
517 gGlyPheValProPheValProGlySerProThrGluArgAsn 531
1847 CGGGTTCGTGCGCTTCTCCCTGGATCACCAACCGAGCAGAGC 1889

seq_name: gb_pl1:AC009465

seq_documentation_block:
LOCUS AC009465 93234 bp DNA PLN 12-OCT-2000
DEFINITION Arabidopsis thaliana chromosome 3 BAC T9J14 genomic sequence,
complete sequence.
ACCESSION AC009465
VERSION AC009465.5 GI:5932531
KEYWORDS HTG.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryote; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsis.
Lin,X., Kaul,S., Town,C.D., Benito,M., Creasy,T.H., Haas,B.J.,
Wu,D., Maiti,R., Ronning,C.M., Koo,H., Fujii,C.Y., Ulterback,T.R.,
Barnstead,M.E., Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.
Arabidopsis thaliana chromosome 3 BAC T9J14 genomic sequence
Unpublished
REFERENCE 2 (bases 1 to 93234)
AUTHORS Lin,X. and Kaul,S.
```

Direct Submission
Submitted (24-AUG-1999) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, xlin@tigr.org
3 (bases 1 to 93234)
Lin,X.
Direct Submission
Submitted (29-SEP-1999) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
4 (bases 1 to 93234)
Town,C.D. and Kaul,S.
Direct Submission
Submitted (12-SEP-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
5 (bases 1 to 93234)
Town,C.D. and Kaul,S.
Direct Submission
Submitted (05-OCT-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
6 (bases 1 to 93234)
Town,C.D. and Kaul,S.
Direct Submission
Submitted (12-OCT-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
On Sep 29, 1999 this sequence version replaced gi:5809709.
Address all correspondence to:at@tigr.org

BAC clone T9J14 is from Arabidopsis thaliana chromosome 3
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of several methods: Gene
prediction programs including Genscan+ (Chris Burge,
http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHM (Mark Borodovsky,
http://genemark.biology.gatech.edu/GeneMark/), GlimmerA (a variant
of GlimmerM, see Mihaela Pertea,
http://www.tigr.org/softlab/glimmerm.htm/glimmerm.html, and
GeneSplicer (Mihaela Pertea and Steven Salzberg, contact
mpertea@tigr.org), searches of the complete sequence against a
peptide database and the plant EST database at TIGR
(http://www.tigr.org/tdb/tgi.shtml). Annotated genes are named to
indicate the level of evidence for their annotation. Genes with
similarity to other proteins are named after the database hits.
Genes without significant peptide similarity but with EST
similarity are named as unknown proteins. Genes without protein
or EST similarity, that are predicted by more than two gene
prediction programs over most of their length are annotated as
hypothetical proteins. Genes encoding tRNAs are predicted by
tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).
Simple repeats are identified by repeatmasker (Arian Smit,
http://ftp.genome.washington.edu/RM/RepeatMasker.html).

Location/Qualifiers
1. 93234
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="III"
/map="mil72"
/clone="T9J14"
complement(1..945)
/note="overlap with BAC clone T12H1 [AC009177:1..945]."
join(<244..358,454..623,755..856,1190..1293,1386..1491,
1802..2043,2141..2639,3121..3351)
/gene="T9J14.1"
244..3351
/genes="T9J14.1"
/note="predicted by genefinder"
join(244..358,454..623,755..856,1190..1293,1386..1491,
1802..2043,2141..2639,3121..3351)
/gene="T9J14.1"
/codon_start=1
/product="unknown protein; 244-3351"
/protein_id="AAD56987.1"
/db_xref="GI:5932532"
/translation="MDSQQLALDCIVATLFDGSNEFGSSSEVHALRGIFEGLLQQL"

FEATURES
source

misc_feature
mRNA
gene
CDS

```
LSLKNWPELMKVHVHYLDAMGPFELKYPPDAVAGSLINKLFELLTSLPHVVKDPATPTS
RAARLIQCTSTRIATAKAESKSVLPHMGIAIDTMGYLADKEGTLRLRGEHNLSEAFVLVA
SSAGAQQOQEWLAWLEPLLSQOWPEMNQVNSLDPMLCLLRVLHLSLWSPVSQFPLPP
EPRAAKMTWDAERYSLGSEANPKLSKGVSVVADGSFECTKEGQAEASESDIRNMLKGI
RDCGNVLGLSTTIGTEFFKCLDANYVAMALMENLQSMFEFRHILRFIHTFTIYLVKSC
PADMWESVLWGLHPLFTHCOOLSSAWPGLLQEGRAKVPDLFGIQSODMKLWEMEE
KILLDTREIATLFSTMASPGLTGVPLVLESHGVGRVDMSTLDTLHAFRSNVMVLS
LPCLTPNDLHAFEAETAKTSPKQKQLMRSLLLIGTGNLKAALAAQKSONVITNVTG
NKVSLHAHDISVFF"
Join(<8161, 8331, 8463, 8581, 9011, 9108, 9205, 9264,
9340, 9386, 9475, 9669, 9764, 9810, 9892, 9958, 10061, 10161,
10253, 10354, 10446, 10584, 10669, 10731, 10832, 10915,
11142, 11489)
/gene="T9J14.2"
8161, 11489
/gene="T9J14.2"
/translation="similar to sodium proton exchanger Nhx1 GB:AAD16946
[Arabidopsis thaliana]"
Join(8161, 8331, 8463, 8581, 9011, 9108, 9205, 9264,
9340, 9386, 9475, 9669, 9764, 9810, 9892, 9958, 10061, 10161,
10253, 10354, 10446, 10584, 10669, 10731, 10832, 10915,
11142, 11489)
/gene="T9J14.2"
/translation="similar to sodium proton exchanger; 8161-11489"
/product="putative sodium proton exchanger; 8161-11489"
/protein_id="AAD56988.1"
/db_xref="GI:5932533"
/translation="MTMFASLTSMKLSVSTSDHASVVSNLFLVALLCACIVIGHLLEE
NRWNESITALLIGLGTGVVILLISRGKNSHLLVSEDLFFIYLLPPIFNAGFOVKK
KQFFNFVIMAFGAIGTVVSCITISLGAIQFFKKLDIGTDFGLDELAIGAFAATDS
VCTGLAQDETPLLSYLVFGVVDNATSVLFAQSFDLTHLNEAEPFQJGNPF
YLFLLSTGVATGLISAYVKKLYFGRHSTDREVALMMLMAYLSYMAELFLSGIL
TVFFGIVMSHYTNWNTVSSRITTKHAFATLSLAETFFIFYVGMALDIEKWRVYS
DSPGTSVAVSLTGLVLMGRAAVPEPLSFLSNLAKKHQSEKISIKQOVVLWAGLAR
GAVSALAYNFTSRSHLELGNALMITSITVCLFSTWVFGMLTKELIRLMPHQKA
TYTSTSDLSHSDTPKSIHPLLDGEQDSFELPGSHQDVPKPSNLGFLMRPRTVHY
YWRQDFDAFMRPVPFGRGFVFPVPGSPTESSHLSKP"
Join(<12067, 12114, 12560, 12676, 12757, 12879,
12957, 13082)
/gene="T9J14.3"
12067, 13082
/gene="T9J14.3"
/translation="T9J14.3"
/translation="identical to GB:P11829 [Arabidopsis thaliana]"
Join(12067, 12114, 12560, 12676, 12757, 12879, 12957, 13082)
/gene="T9J14.3"
/codon_start=1
/product="acyl carrier protein 1 precursor (ACP);
12067-13082"
/protein_id="AAD56989.1"
/db_xref="GI:5932534"
/translation="MATQFSASVSIQTSCLATTRISFOKPALISNHQKTNLSFNLRRS
IPSRRLSVSCAAKQETIEKVAIVKKQLSLTPDKKVAETKFDLIGADSLDTVEIVMG
LEEFNIOAEKAKOIKIATVQAAELIEELINEK"
complement(join(<13275, 13560, 13812, 13893, 13995, 14097,
14170, 14365, 14560, 14751, 15058, 15342))
/gene="T9J14.4"
complement(13275, 15342)
/gene="T9J14.4"
/translation="T9J14.4"
/translation="similar to GB:AAB61079 (contains weak similarity to
the SAPB protein (TR:E236624)) [Arabidopsis thaliana]"
complement(join(13483, 13560, 13812, 13893, 13995, 14097,
14170, 14365, 14560, 14751, 15058, 15342))
/gene="T9J14.4"
/codon_start=1
/product="unknown protein; 15342-13483"
/protein_id="AAD56990.1"
/db_xref="GI:5932535"
/translation="NRVLSEIAESPFVISRLSPDSTATGGFIGNGVGKCHGLHNTVL
VIASLILVAYLAYEAKSLSKLSNRRSYIMIAYYGFLWLVLNLLANWCCQAWBCTPG
KEVINLLTLFTTSGMLSVLAFPLQGNVYAGAEALTRTFILISGLVIGLDLLKA
YLFYCGVPLFDNNEHLHKPKWGLWVHKLLLAGIYGMIFEMVNSKWRRLPCEYID
HSYPPYAPAFYKYITVMLALNGLSLFACALTANGAHFGLWLYGITSVCYHAFLPLL
YVTFADLFFQEDLNLNYYSEMKGDFDADME"
```

```
mrna
Join(16215, 16399, 16873, 16990, 17073, 17295, 17473, 17732)
/gene="T9J14.5"
16215, 17732
/gene="T9J14.5"
/translation="MGREVAESCVDTMLMEMVAMYSGRFYANKPELAARRIEAIGQV
CHOLSRYTMRERPRFSDHLEAIFICKDFWSEVFKQIDNLTNHRGTFLQDNKFRW
LSRSIDPSSSENETQDPSTPGESKAAQAVSMVLYFPGGIIRGVLSNLGIPCAVSADIS
SLPTCSFVIRVKA"
19191, 19378
/translation="unknown protein; 16248-17501"
/protein_id="AAD56991.1"
/db_xref="GI:5932536"
/translation="MGREVAESCVDTMLMEMVAMYSGRFYANKPELAARRIEAIGQV
CHOLSRYTMRERPRFSDHLEAIFICKDFWSEVFKQIDNLTNHRGTFLQDNKFRW
LSRSIDPSSSENETQDPSTPGESKAAQAVSMVLYFPGGIIRGVLSNLGIPCAVSADIS
SLPTCSFVIRVKA"
19191, 19378
/translation="Emigrant MYTE-like element from T517 [17254 to
17779] 526 nt"
<20266, 20949
/gene="T9J14.6"
20266, 20949
/gene="T9J14.6"
/translation="predicted by genefinder"
20266, 20949
/gene="T9J14.6"
/product="hypothetical protein; 20266-20949"
/protein_id="AAD56992.1"
/db_xref="GI:5932537"
/translation="MEKLRADAAEIMICAGQLKGLNLHLMRGMTEANIEKSSRELDLKEK
ELQILSSDILQGFSAPEAKSEVDDKLKLVCECTEELSRKNLLTAVKLDLSLRVQREL
ELKDNLQVQMAELKRRYSEARHVQRKREMETATKKKLSMTVDQIQESGKQLEK
KSREVELKQKEIEBKGLDLVKQVKAWEKLIQLSKLVDDDCCTTELSPRKDQVDS
NNIYVVELG"
<22764, 26261
alignment_scores:
Quality: 1924.00 Length: 1110
Ratio: 3.693 Gaps: 22
Percent Similarity: 46.937 Percent Identity: 42.613
alignment_block:
US-09-271-584A-2 x AC009465
Align seg 1/1 to: AC009465 from: 1 to: 93234
1 MetLeuAspSerLeuValSerIleuProSerLeuSerThrSerAspHi 17
8167 ATGTTGCGCTCTTTAACTCTTAAATGCTATCGGTGTCACATTCGATCA 8216
17 sAlaSerValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysI 34
8217 CGCATCTGTCGTTTCACITAACTCTTTGTTGTCCTTCTATGTGCTTGA 8266
34 leValLeuGlyHisLeuLeuGluAsnArgTrpMetAsnGluSerIle 50
8267 TCGTCATTTGCCATCTTTTGGAGGAGAAATCGATGATGAACCAATCCATC 8316
51 ThrAlaLeuLeu
8317 ACTGCTTTATTGATTGTGAAGTTTCCATTATTTTGGCTTGAATGGTAGCTA 8366
54
8367 AAGAGGAGAGCGATAAGTTTTACATTTTGGATATTATCATATCAATGTT 8416
55
8417 TCACITTCATTTTCTTTACTGACTTTTAAAGCAATGGGTTGCGAGGGC 8466
57 euGlyThrGlyValThrIleLeuLeuIleSerLeuLysGlyLysSerHis 73
8467 TTGGCACTGGTCTGTCATATTTGTTGATTAGTAGAGGAGGAAAACATCAT 8516
```

74 LeuLeuValPheSerGluAspLeuPheIleTyrLeuLeuProIle 90
|||||
8517 CTCCTGGCTTCTAGTGAAGATCTCTCTTATATATCTTTGGCACCAT 8566
90 eilePheAsnAlaGly..... 95
8567 AATATTCAATGCAGGGAAGTTTCTACATCCGAGCTGATCATTTGATGAT 8616
95 95
8617 AGATAGTTGCATATTTTGTGCTCTTCTCTATTTGATACATCTTGTGTCT 8666
95 95
8667 AGTAATATGTTTGGTTGGTTGTACTACAGAGGGGAAAAATCAGTTTA 8716
95 95
8717 TTATCAGAGATCTTACTGGGAAATGTTATACATACATTTTATAAC 8766
95 95
8767 ATATATGCATTAATGCTGATCTTATGAATATGTTTCAAAATACAGCTTG 8816
95 95
8817 GAAGTACTTGTGATTTGCTACTGCAGAAATAGGAGTGGGTATATGTGAT 8866
95 95
8867 GGTCATTTTGTGATTTATCTAAAGTGGTCTTCTATGCTTTGTGCG 8916
95 95
8917 TAAATGCAGACATGGCCAGCTTTCTTGTTCACACTATGTTTCTCAG 8966
96 PheG1 97
8967 ATCTTCTCTGAACCAATAATTTTGTGTTGCTGTTATTTAGGTTTCA 9016
97 nValLysLysLysGlnPhePheArgAsnPheValThrIleMetLeuPheG 114
|||||
9017 AGTAAAAAAGACGAGCTTTTCGAAATTTTGTAACTATATATGCTTTG 9066
114 lYAlaValGlyThrIleIleSerCysThrIleIleSerLeuGlyVal... 129
|||||
9067 GCGCCATTGGGACCGTAGTTTCTTGACCAATAATATCTAGGTATCTTA 9116
129 129
9117 CTCCCTCTCCCTGTAATGATTCGTGTGAAGCTTACTCCTGATTAGTA 9166
130 ThrGlnP 132
9167 ATTGGGATCAATAATATCTTGTGATGTTTGAACAGGTGCAATTCAGT 9216
132 hePheLysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeu... 147
|||||
9217 TCTTTAAGAAATAGACATTGGGACCTTTGACTTGGGCGATTTCTCTGGT 9266
147 147
9267 AATGTGTTAAGGAGTTCTTTTATCGATTTTGTCTAGAGACTTGAT 9316
148 AlaIleGlyAlaIlePheAlaAlaThrA 157
9317 TTTTCTTTTGTCTTACTTAATCAAGCAATCGGCGCATATTTGCTGCAACCG 9366
157 spSerValCysThrLeu.Gln..... 163
9367 ACTCTGTATGCACATACAGGTAATGTGATTCGCCAATATACAGAAAGTAC 9416

163 163
9417 ATACAGGCCAGCAATCAAAGTCTAATTAACAAAAACTCTTGTTCCTTTGG 9466
164ValLeuAsnGlnAspGluThrProLeuLeuTyrSerLeuVal 177
|||||
9467 TGCATACAGGTTCTCAATCAAGATGAGACACCTTTGCTTTACAGTCTGTGA 9516
178 PheGlyGluGlyValValAsnAspAlaThrSerValValValPheAsnAl 194
|||||
9517 TTTGGAGAGGGCTTGTGAATGATGCCACATCTGTTGTCTCTCAATGC 9566
194 aileGlnSerPheAspLeuThrHisLeuAsnHisGluAlaAlaPheHisL 211
|||||
9567 TATTTCAGAGTTTGTGACCTCACCCACTTAACCATGAAGCAGCTTTTCAAT 9616
211 euLeuGlyAsnPheLeuLeuPheLeuLeuSerThrLeuLeuGlyAla 227
|||||
9617 TTTCTTGGGAACCTTTTATCTGTTTCTTGTGAGCACCGGACTTGGTGTG 9666
227 227
9667 GCAGTAAGTCCATTTTCTCTCTACTGCTCTTTCGCTGGCTACTACT 9716
228Ala.Thr 229
9717 TCCAGGTTCCCTCTGTTAACTACAATTTTTTGTATTATCATCTTGCAGACT 9766
230 GlyLeuIleSerAlaTyrValIleLysLysLeuTyrPheGly..... 243
|||||
9767 GGTCTGATAAGTCTTATGTCATCAAGAACTGTATTTTGAAGGTAGAT 9816
243 243
9817 AATTGGCTTTAGCTCATATGTGACTTTTATCTCTATTAGTCTTCTGTTATG 9866
244ArgHisSerThrAspArgGluValAla 252
9867 TAGTCTAATTTGGGTATCTTTCTTGGGCACTCGACTGATCGAGAAAGTTGCC 9916
253 LeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGlu..... 266
|||||
9917 CTCATGATGTTATGGCTTATCTTTTCATATATGCTTCTGAGGTTAGTTT 9966
266 266
9967 TGTCAAAATTTTATACCACTGTCTAATAATTTAGGCATGGAGACAATCAGTTG 10016
267Leuphe 268
10017 TATCTTAAGAAATCCTTTATTGATAATATATTGTTGGTTCGACGTATTC 10066
269 AspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMetSerHi 285
|||||
10067 GCCTTGTAGTGTCTATCCTAACTGTATTTTCTGTGGGATTGTGATGCCCA 10116
285 sTyrThrTriPheHisAsnValThrGluSerSerArgIleThrThr..... 299
|||||
10117 TTACACTTGGCAAAATGTCCCGAGAGCTCAAGAAATTACTACCAAGTAAT 10166
299 299
10167 ATGACCTCCATTTTCTCTTGAACAAATTCAAATGTTTCCCATCTTCAA 10216
300Lys.HisThrPheAlat 305
10217 TACTTACTTACTTCTATGCATTTTACGATGAGACAGCAGCTTCCTTTCCTA 10266
305 hrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrValGlyMetAsp 321
|||||
10267 CTTTGTGCTTCTCGCTGAGACTTTTATTTTCTCTACGTTGGAATGAT 10316
322 AlaLeuAspIleAspLysTriPArg.Ser..... 330

```
|||||.....|
10317 GCATTGGACATAGAGAAATGGAGATTTCGTGAGTCACAGGTATGTCCTC 10366
330 .....
10367 AGTAGAAACATAAACTCTCTTTTGTTCATGGCTATATACATAATCTTC 10416
331 .....ValSerAspThrProGlyThrSerIleAlaVal 341
10417 TTATGTTTCCATTGTTGGTTTGGACAGCCCGGGACATCAGTTGCAGT 10466
341 lserSerIleLeuMetGlyLeuValMetValGlyArgAlaAlaPheValP 358
10467 GAGCTCAATCTTAATGGGCTAGTCATGCTTGGAGAGCAGCTTTGTCT 10516
358 heProLeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGluLys 374
10517 TTCCCTCTTCTCTTATCAAACTTAGCCAAAGACATCAGAGCGAGAAA 10566
375 lIeAsnPheAsnMetGln..... 380
10567 ATCAGCATCAAGCAGCAAGTACGAACAATCTACTTTGTCTATTGAATCA 10616
380 .....
10617 AATTTGCTGAGATTGTTGATCTACTTAATCTCTCTCCTTTGCTGT 10666
381 ..ValValIleTrpSerGlyLeuMetArgGlyAlaValSerMetAla 396
10667 AGTTGTGATCTGGTGGCTGTCTTAATCAGAGGTGCTGTATCTATGGCT 10716
397 LeuAlaTyrAsn.Lys..... 401
10717 CTTGCTCAATAAGGTAAATTTGAAAGTTAATGGATAGTAAACAGCGAC 10766
401 .....
10767 ATGGTGATTAGTAAAAATCTACGACATAATTTAACTTTTCTTCTCTG 10816
402 .....PheThrArgAlaGlyHisThrAspValArgGlyAs 413
10817 TGCTCTCTAATCGAGTTTACAGATCAGGGCACAGAAATGCGGGGAA 10866
413 nAlaIleMetIleThrSerThrIleThrValCysLeuPheSerThr..... 428
10867 TGCAATCATGATTACCAAGTACAAATACCGTCTGTCTTTTAGCACCATGG 10916
428 .....
10917 TAAGATATCTCCCTCTCAGTTTGGACTACTAAATGTATTTTCTCTACAG 10966
428 .....
10967 AGATGTTGTCATAGGATAAACCAATGCTGAGCGGTGAGAGGATTTGC 11016
428 .....
11017 ATTTTCTATATCCGGAATACTGTTCCTCTATTGATTGTTGTTATTCAA 11066
428 .....
11067 TGTTTCTGGATGCTAACACTCAACCTTGATACCGGACATTTAAATAT 11116
429 .....Val.ValPheGlyMetLeuThrLysProL 438
11117 GTTTTTTTTTTCCAGGGCTTTTCAGGTGTTTGGTATGCTTAACAAACAC 11166
438 euIleSerTyrLeuLeuProHisGlnAsnAla.....ThrThrSer 451
11167 TGATTAGATACCTTAATGCCACCAAAAGCCACCACTAGTACCAGAGT 11216
452 MetLeuSerAspAsnThrProLysSerIleHisIleProLeuLeuAs 468
|||||.....|
```

```
11217 ATGTTATCGGAGCATAGCACTCCGAAATCAATCCACATTCCTCCTCGA 11266
468 p.....GlnAspSerPheIleGluProSerGlyAsnHisAsnValP 482
11267 TGTGTGAACAGCTAGATTCAATTTGAGTTACCTGGGAGCCACCAGGACGTGC 11316
482 roArgProAspSerIleArgGlyPheLeuThrArgProThrArgThrVal 498
11317 CAGGACCAACAGCCTTCGAGGTTCTCTCATGCGCCCCACACGACTGTC 11366
499 HisTyrTyrTrpArgGlnPheAspSerPheMetArgProValPheGl 515
11367 CACTATTACTGGAGACAGTTTGATGATGCTTCATCGCTCTGTGTTGG 11416
515 yGlyArgGlyPheValProPheValProGlySerProThrGluArgAsn 532
11417 TGTGTCGGGATTGTTCCCTTTGTCTGCTGCTCGACTGAGAGGCA 11466
532 roProAspLeuSerLys 537
11467 GCCATGATCTTAGTAAA 11483
```

seq_name: gb_p12:ATAC011623

seq_documentation_block:

LOCUS ATAC011623 86022 bp DNA PLN 16-NOV-1999
DEFINITION Arabidopsis thaliana chromosome III BAC F24P17 genomic sequence,
complete sequence.
ACCESSION AC011623
VERSION AC011623.3 GI:6437540
KEYWORDS HTG.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1 (bases 1 to 86022)
AUTHORS Lin,X., Kaul,S., Town,C.D., Benito,M., Creasy,T.H., Haas,B.,
Ronning,C.M., Koo,H., Fujii,C.Y., Utterback,T.R., Barnstead,M.E.,
Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.
Arabidopsis thaliana chromosome III BAC F24P17 genomic sequence

REFERENCE 2 (bases 1 to 86022)
AUTHORS Lin,X. and Kaul,S.
TITLE Direct Submission

REFERENCE 3 (bases 1 to 86022)
AUTHORS Lin,X.
TITLE Direct Submission
JOURNAL Submitted (08-OCT-1999) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA, xlin@tigr.org

COMMENT

On Nov 16, 1999 this sequence version replaced gi:6091707.
Address all correspondence to:
Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr.
Rockville, MD 20850, USA
e-mail: xlin@tigr.org

BAC clone F24P17 is from Arabidopsis chromosome III and is near the
molecular marker mi403.
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of three methods: Gene
prediction programs including GRAIL (available by anonymous ftp
from arthur.epm.ornl.gov), Genefinder (Phil Green, University of
Washington), GenScan (Chris Burge,
http://www.cbs.dtu.dk/netgene/cbsnetgene.html), and NetPlantGene
complete sequence against a peptide database and the Arabidopsis
EST database at TIGR (http://www.tigr.org/tdb/at/at.html).
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named

after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are identified by RepeatMasker (Arjan Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Regions of genomic sequence that are not annotated as genes but have predicted exons by GRAIL are annotated as misc features.

FEATURES

source

```

1. .86022
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="III"
/map="mi403"
/clone="F24P17"
join(<283. .529,638. .818,916. .1063,1101. .1142,1160. .1376,
1511. .1822)
/gene="F24P17.1"
<283. .1822
/gene="F24P17.1"
/note="similar to pyruvate dehydrogenase kinase
GB:AAC97601, [Arabidopsis thaliana]"
join(<283. .529,638. .818,916. .1063,1101. .1142,1160. .1376,
1511. .1569)
/gene="F24P17.1"
/codon_start=1
/product="putative pyruvate dehydrogenase kinase, 5'
partial"
/protein_id="AAF08568.1"
/db_xref="GI:6437541"
/translation="VRDWYLFSDMRAPFEIKSGDEKDFQTMKAVKVRHNNVPM
MALGVNQKMGNSDELDFHOFDRYLSRIGRMILQGHVELHNPPLHTVGYIH
TKMSPMEYARNASDEADISCFREYSGAPEINIGDPSFTFPYVPHLMMYELVKN
LRAVQERVDRAVPPIRIIVADGIEDVTIKPFRSLHLRFPDIIVVSDGEGGLARSG
LPRIITYLSTARNPLEDVLGIADVPVIMAGVGYGLPISRLYARVFGDGLQIISME
CYGTDAYLHLSRLGDSQEPPLP"
complement(2022. .2051)
/rpt_family="POLY_A"
complement(2053. .2372)
/note="exon predicted by xgrail, quality
excellent_shadowexon"
2057. .2160
/note="exon predicted by xgrail, quality excellent"
2287. .2363
/note="exon predicted by xgrail, quality excellent"
join(<2630. .3712,3813. .4007,4256. .4324,4406. .4675,
4767. .5117,5215. .5492,5579. .>6599)
/gene="F24P17.2"
<2630. .>6599
/gene="F24P17.2"
/note="similar to RNA helicase (DRH1) GB:BAA28347
[Arabidopsis thaliana]; Pfam HMM hits: WW/rsp5/wmp domain
containing proteins, DEAD and DEAH box helicases,
Helicases conserved C-terminal domain"
join(2630. .3712,3813. .4007,4256. .4324,4406. .4675,
4767. .5117,5215. .5492,5579. .6599)
/gene="F24P17.2"
/codon_start=1
/product="putative RNA helicase"
/protein_id="AAF08584.1"
/db_xref="GI:6437557"
/translation="MATEDTPASAGRPYAPEDTLPQPKGLIDGSTGILYVNPPT
NVTQYRPSAPPPIHSATTPKLAQIPVPSGGHQAHEQAKPVGHVSQQHGFQQOQPO
FPSQVRPQMQRHQAQMQPQQSQQFPQQSQSMVPHPHCHPSVQTYQPTTQQOQOG
MONSHSQPOOLSHOYAHSOOHYMGFRPHMOTGLONSHOTPGQPHGQFPQSQEYN
SLAPKREDEPHGCKTKTGSQPHLPNSRSPSONTHEEANAASOKTNANLMAQKNG
POANNAVTOFOQGANLIHOOLGRAPNQMDQTMHLHOKSHVSPSQNNTYENHQSRR
GNDSTVNNRMEVPVRGAQPLHPAAMPKDIRISGGPPTNADPAMQGTGHGYHAGPAF
PNKSLVRPHFVTSDPVPLHSPVEIYRQHEVTTTGTENIPAYITTFESSGLPPEILREL

```

```

LSAGFPSPPTIOAQTFIALQSRDIAVAIAKTGSGKTLGYLIPAFILLRHCNRDSNCP
TVLILATPRLAQTIODEALRFGRSSRISCTCLYGGAKPGQPKQLKEGRADIIVATPG
RLNDILEMKMIDFQOVNGLRVDLEAANKAITQYVVPQMEKERRLEQILRSQERGSK
EVRIASDLNVNPNVGNIRVDELAANKAITQYVVPQMEKERRLEQILRSQERGSK
VIIFCSTKRLCDHLARSVGRFAGVITHGDKTQGERDVLNQFRSGKSCVLIAIDVAA
RGLDIKDIRVINYDFPTGVEDYVHRIGRTGRAGATGVAFTTEODWKVAPDLIKVL
ESGNGQVPPQVDIRIAMRGGGGGPGYSQDRRGVMNRFDSGGGTWDSGGGFGGCGG
FGGREGGFGGREGGFGGREGGFGGREGGFGGREGGFGGREGGFGGREGGFGGREGG
RGFGFRGNMNMESRGFGSGRGFGRGFGRGFGRGFGRGFGRGFGRGFGRGFGRGFG
KRSRGSYSRSRSRSRSRSRSPRSHRDRGHNRSRSYRSRSPVYRERDRAPR
VSGFDIKPPVSVVNVLDNMAAAAIENVVPTSLSEROGNGVSEVEAALVRPVVDEEP
5843. .5962
/rpt_family="(TGG)n"
complement(7077. .>8090)
/gene="F24P17.3"
complement(7077. .>8090)
/gene="F24P17.3"
/note="similar to hypothetical proteins GB:CAB16818,
GB:AAD55277"
complement(7254. .8090)
/gene="F24P17.3"
/note="unknown protein"
/codon_start=1
/protein_id="AAF08569.1"
/db_xref="GI:6437542"
/translation="MASISLTYLWLVNHPYISNFTWIEGETIGSTVFFVSVVSVYL
SATFLRSADISLPSRPIKLTAVHSLILCLLSLWAVGCTLSITSSHASDDPA
RFLHAICFPYDVKPNPFLFWAQVFLSKILEFGDTLILILKTSIQRLHVFYHAT
VVMVYCLWLTRQSMFPALVSTVHVIMYGYFLCAVGRPKWRLVTDQOIVQFV
FSGLSGWMLREHLFGSGCTGNGWCENAFNASLALFNSFNHKNVKKPTREDGKK
SD"
8262. .8362
/note="exon predicted by xgrail, quality
marginal_shadowexon"
complement(8375. .8450)
/notes="exon predicted by xgrail, quality marginal"
complement(<8719. .>9615)
/gene="F24P17.4"
complement(<8719. .>9615)
/gene="F24P17.4"
/notes="similar to hypothetical proteins GB:CAB16818,
GB:AAD55277"
complement(8719. .9615)
/gene="F24P17.4"
/notes="unknown protein"
/codon_start=1
/protein_id="AAF08570.1"
/db_xref="GI:6437543"
/translation="MASVSTLYLWLVHHPYIANFTWTGSETLGSVTFVFFVSVLYL
SATFLRYTVDSLPTLGPRIKLTAVHSLILCLLSLWAVGCTLSLSSDDPKARLF
DAVCFPLDVKPGPLFFWQAQVFLSKILEFVDTLLILNKSIQRLSFLHVVHATYVI
LCYLWLTRQSMFPVGLVNSTVHVIMYGYFLCAIGSRPKWRLVTFNQVQAFGM
GLGAAMWLPEHYFGSGCAGIWTYVGVFTASLLALFYNFHNSKNYEKTTTPLYKIES
FIFIHGERWANKAITLFSKND"
9798. .9853
/note="exon predicted by xgrail, quality good_shadowexon"
complement(10146. .11457)
/gene="F24P17.5"
/notes="similar to putative protein GB:CAB45783"
/pseudo
complement(12632. .12715)
/notes="exon predicted by xgrail, quality
marginal_shadowexon"
join(<12993. .13102,13175. .13257,13357. .13549,13667. .13805,
13900. .14065,14182. .14339,14418. .14612,14700. .14795,
14879. .15013,15093. .15400,15493. .15763,15855. .>16211)
/gene="F24P17.6"
<12993. .>16211
/gene="F24P17.6"
/notes="similar to putative Anion exchanger family members:
GB:AAD39673, GB:AAD55295 [Arabidopsis thaliana]"
join(12993. .13102,13175. .13257,13357. .13549,13667. .13805,

```

13900..14065,14182..14339,14418..14612,14700..14795,
14879..15013,15093..15400,15493..15763,15855..16211)
/gene="F24P17.6"
/note="unknown protein"
/codon_start=1
/protein_id="AAF08571.1"

alignment_scores:
Quality: 1492.50 Length: 884
Ratio: 3.116 Gaps: 22
Percent Similarity: 54.186 Percent Identity: 42.308

alignment_block:

US-09-271-584A-2 x ATAC011623 ..

Align seg 1/1 to: ATAC011623 from: 1 to: 86022

2 LeuAspSerLeuValSerLysLeuProSerLeuSerThrSerAspHisAl 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
58793 TTAAGCAATGCTGGAAAAACAGAGGCTCTTTTCGCTTCTGATCATGC 58842
18 aSerValValAlaLeuAsnLeuPheValAlaLeuLeuCysAlaCysIleV 35
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
58843 ATCGGTGCTGCCATGAATTTGTTTCGAGCTTTTTCGTTTGCTGTCATCG 58892
35 alLeuGlyHisLeuLeuGluAsnArgTrpMetAsnGluSerIleThr 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
58893 TGTCTTGTCACCTTCTTGAGGAGACTCGGTGGATGAACGAGTCAATCACT 58942
52 AlalaLeuIle..... 55
|||||:|||||
58943 GCTCTTATCATTTGAAGCTTCGTTTTTTTTTTTCCCAAAATCTTTAAAC 58992
56GlyLeuG 58
|||||
58993 ACATTGCTTACAACATCTCTAAATTTGGATTACGCAAAATAGGTTTCGT 59042
58 lyThrGlyValThrIleLeuLeuIleSerLysGlyLysSerSerHisLeu 74
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
59043 GTACTGGGATTGTGATCTTGCTTATAAGTGGAGGCAAAAGCTCAAGATT 59092
75 LeuValPheSerGluAspLeuPhePheIleIleIleLeuLeuProIleIle 91
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
59093 CTGTGTTTGTAGTGAAGATCTCTTCTTATTTATCTTCTCCACCAATTAT 59142
91 ePheAsnAlaGly..... 95
|||||:|||||
59143 ATTCACCGCAGGTAAGTCTCTTTTATATAGAAATTCATGAAAAATTGAA 59192
96PheGlnValLysLys 100
|||||:|||||
59193 AGTCGGAATTGAGTAACACATTTTTTTTGTACAGGTTTCAGGTTAAGAAG 59242
101 LysGlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaValG 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
59243 AAGCAATTTTTTCGCACTTCATGACCAATATGTTATTTGGTGCTATTGG 59292
117 yThrIleIleSerCysThrIleIleSer..... 126
|||||:|||||:|||||:|||||
59293 AACTCTCATTTTCATTGTTATCTCATCTTGGTAAAAAACCAATCTCAA 59342
126 126
59343 TCAACTCCACATTTCAATATCATGAGTCATATATTTTGTATGATGTTT 59392
127LeuGlyValThrGlnPhePheLysLysLe 136
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
59393 CCGGTTTTGTTTGTATGCTAGGTGCTAAACATCTTTTCGAGAAAAAT 59442
136 uAspIleGlyThrPheAspLeuGlyAspTyr..... 146
|||||:|||||:|||||:|||||
59443 GAATATCGGTGATCTTACCATTGCGGACTATCTAGGTACCTGAAACAAA 59492

146 146
59493 GAAACCCCTTTTCAAGTCTCTTATTCATTAAAGTCTCATACTGTTTAAATCT 59542
147Leu..AlaIleGlyAlaIlePheAlaAlaAlaThrAspSer 158
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
59543 GTACTGATGATTTTGCAGCCATTGGAGCAATATCTCTGCTACAGACTCT 59592
159 ValCysThrLeu.Gln..... 163
|||||:|||||:|||||:|||||
59593 GTTTCACCTTCGCAAGTAGAAGACTTTTTCGAAACATTCAGAGATATT 59642
163 163
59643 ATTTCTACTACATTTTCTAATCAAAAATTAACCTGACGCTTTGTTAAATC 59692
164ValLeuAsnGlnAspGluThrProLeuLeuTyrSerLeuValPheG 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
59693 TAGGTGCTTAAATCAAGACGAGACACTCTCTTGTACAGTCTTGTCTTTGG 59742
179 yGluGlyValValAsnAspAlaThrSerValValValPheAsnAlaIleG 196
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
59743 AGAAGGTGTAGTGAAGATGCCACATCGTCTGCTCTTCAATGCAATAC 59792
196 InSerPheAspLeuThrHisLeuAsnHisGluAlaAlaPheHisLeuLeu 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
59793 AGAGATTCGACCTCAAAATATCAATTCAGCCATAGCTTTGGAGTTTGT 59842
213 GlyAsnPheLeuTyrLeuPheLeuLeuSerThrLeuLeuGlyAlaAla... 228
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
59843 GGAACATTTTTCCTCTCTTATCTTAAGCACAGCACTGGTGTGTCAGT 59892
228 228
59893 AAGTTTCTTCTCGCTCTCAAAATTTCTAAACCGGAAAAATTTCTCAC 59942
229ThrGlyLeuIleSerAlaTyrValIleIleLysLysLeu 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
59943 TCGTTTCTTTTAGGCTGGATTGCTCAGTCTTTTGTATCAAGAAGCTC 59992
241 TyrPheGly..... 243
|||||:|||||
59993 TATATAGGAGGTATATTCGTATCTACTCATGATGTTAAAAATGTGATTTA 60042
244A 244
60043 CCCTTCTTTTAAATATCTTTTGTCTTTTGTCTTTTGTCTTTGTAACAT 60092
244 rg..HisSerThrAspArgGluValAlaLeuMetMetLeuMetAlaTyrL 260
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
60093 CAGGCACCTCTACTGATCGTGAAGTTGCACCTTATGATGCTATTGGCTTACT 60142
260 euSerTyrMetLeuAla.Glu..... 266
|||||:|||||:|||||:|||||
60143 TATCATATATGTTGGCAGAGTAATAAGAAAAATTTGAACGTTTGGTT 60192
267L 267
60193 ACATTTTTCACAGAAAAATGAAACATGATACATATATCTCTTTCGAGC 60242
267 eupheAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMet 283
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
60243 TATTCACCTTGAGCTCTATCTTGAAGTGTCTCTCTGCGGATGTTATG 60292
284 SerHisTyrThrTrpHisAsnValThrGluSerSerArgIleThrThry 300
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
60293 TCTCATATACATGGCAACAATGTTACAGATAAATCCAAGTCACTACAAA 60342
300 s..... 300
60343 GTAAGTTTCTTTTCTTCCCAATAAATAGAAAAATACTTACTTAATTT 60392
301HisThr 302

```
60393 GTTGTGCTTGAATGACTAAACTGATACGTTTCTTGGCGTAGACATCT 60442
|||||
303 PheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrValG1 319
|||||
60443 TTTGCTGCAATGCATTTCTAGCTGAGATTTTATCTCTCTTACCTGG 60492
|||||
319 yMetAspAlaLeuAspIleAspLysTrrp..ArgSerValSerAsp..... 333
|||||
60493 AATGGACGCTCTCGATATCGAGAAATGGACGTTGTACCAACAGGTGAG 60542
333 ..... 333
60543 ACTTAGAGTGTGATAATACATCAATATCTTTGTTTCAFTTTCATAATAT 60592
334 .....ThrProGlyThrSerIleAlaValSerSerIle 344
|||||
60593 TGTTTAACAAATTTTGGCAGCTCTGGTGCAGTGCAGTGGAGTTAGTTCAATA 60642
|||||
345 LeuMetGlyLeuValMetValGlyArgAlaAlaPheValPheProLeuSe 361
|||||
60643 CTCTTGGGCTTATCTCTGGTGGCGCGGCTTCGTTTCCACTTTC 60692
|||||
361 rPheLeuSerAsnLeuAlaLysAsnGlnSerGluLysIleAsnPhEa 378
|||||
60693 CTCTTGTGTCAAATTTAACAAAGCTTTCACCGGATGAGAAATAGACTTAA 60742
|||||
378 snMetGlnValValIleTrrpTrrpSerGlyLeuMetArgGlyAlaValSer 394
|||||
60743 AGAACAAGTAACCAATTTGGTGGGCTGGTCTGATCGGTGGTGCAGTGCA 60792
|||||
395 MetAlaLeuAlaTyrAsnLys..Phe..... 402
|||||
60793 ATGGCTCTTGCTTATAACACAGGTTTGTCCGTGACTACAGAACAACAA 60842
402 ..... 402
60843 GATAAAAAAGTGATCTTTGGGTTTAAATAAGTTCCTCATGACTTAATCTCT 60892
403 .....ThrArgAlaGlyHisThrAspValArgG 412
|||||
60893 TCTCTCTGTATCATATGATGTCACAACTTCAGGACACACCAAGGTTCTTG 60942
|||||
412 LysAlaAlaIleMetIleThrSerThrIleThrValCysLeuPheSerThr 428
|||||
60943 GGAACGCTATCATGATCACCAGTACCATCACTGTGTCTTTTTCAGTACT 60992
428 ..... 428
60993 GTGGTACGTTTCAACACTTTGTTCAAACCGGTTTGTCTTCTTACTGAAGTC 61042
429 .....V 429
61043 AAATCGGGTTTAGCCTATCCACACAGTAACCTAATCTCTACACCTTTTGTG 61092
429 al.ValPheGlyMetLeuThrLysProLeuLeuSerTyrLeuLeuPro.. 444
|||||
61093 CAGGTGTTTGGATTTGCTAACCACCAACCGTTAGTCAACATCTTTCAGCCTTC 61142
445 ....HisGlnAsnAlaThrThrSerMet.....LeuSerAspAspA 457
|||||
61143 ATCAAAACAGTCTCCACACCGCCTCGCATACATCAAGATCTTCTCT 61192
457 snThrProLysSerIleHisIleProLeuLeuAspGlnAspSerPheIle 473
|||||
61193 TCCACGATCCGATCTCCATGAGCGGTGCTCAGTACCACAGGCCAGTCAC 61242
474 GluProSerGlyAsnHisAsnValProArgProAspSerIleArgGlyph 490
|||||
61243 GAATACGACCTCGAACACATGTT.....AGCTTCAGAATGTT 61280
490 eLeuThrArgProThrArgThrValHisTyrTrrpArgGlnPheAspA 507
|||||
```

```
61281 CTGGAAATCTCGTCCAGGGCATTTCATCATTTACTGGAGAAATTCGATA 61330
507 spSerPheMetArgProValPheGlyArgGlyPheValPheValPropheVal 523
|||||
61331 ACGCAGTTTATCGTGGCATATTTGTCGCCGAGCGTTTTCACCATGATT 61380
|||||
524 ProGlySerProThrGluArgAsnProProAspLeuSerLys 537
|||||
61381 CCAGTTCACCCATTGAGAAATAGTTCGCCCAATGGAGTGAA 61422
seq_name: gb_p11:AB033990

seq_documentation_block:
LOCUS AB033990 7545 bp DNA PLN 06-OCT-2000
DEFINITION Ipomoea nil Purple gene for Na+/H+ exchanger, complete cds.
ACCESSION AB033990
VERSION AB033990.1 GI:10716130
KEYWORDS
SOURCE
Ipomoea nil (strain:KKZSK-2) foliage DNA, clone_lib:genomic
library.
ORGANISM
Ipomoea nil
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
1; Solanales; Convolvulaceae; Ipomoea.
REFERENCE
1 (sites)
Fukada-Tanaka,S., Inagaki,Y., Yamaguchi,T., Saito,N. and Iida,S.
Colour-enhancing protein in blue petals
Nature 407, 581 (2000)
REFERENCE
2 (bases 1 to 7545)
Fukada-Tanaka,S., Inagaki,Y., Yamaguchi,T. and Iida,S.
Direct Submission
Submitted (26-Oct-1999) to the DDBJ/EMBL/GenBank databases.
Yoshishige Inagaki, National Institute for Basic Biology, Division
of Gene Expression and Regulation 1; Myodaiji-cho, Okazaki, Aichi
444-8585, Japan (E-mail:yinagaki@nibb.ac.jp, Tel:81-564-55-7684,
Fax:81-564-55-7685)
FEATURES
Location/Qualifiers
1..7545
/organism="Ipomoea nil"
/strain="KKZSK-2"
/db_xref="taxon:35883"
/chromosome="5"
/clone_lib="genomic library"
/tissue_type="foliage"
join(1355..1525,2541..2659,3187..3284,3400..3459,
3558..3604,3698..3892,4014..4060,4180..4246,4362..4462,
4562..4663,5003..5141,5219..5281,5409..5492,5595..5930)
/gene="Purple"
join(1355..1525,2541..2659,3187..3284,3400..3459,
3558..3604,3698..3892,4014..4060,4180..4246,4362..4462,
4562..4663,5003..5141,5219..5281,5409..5492,5595..5930)
/codon_start=1
/product="Na+/H+ exchanger"
/protein_id="BAB16381.1"
/db_xref="GI:10716131"
/translation="MAFGLSLLQNSDLFTSDHASVSMNLFVALLCACTVGLHLEE
ARWVNSITALIGCTGVWILLGGKSHLLVFSDELFFIYLLPPIFNAGFQVKK
KQFFVNFMTIMLFGAIGTILSCISIFGAKIFKHLDFDLDFDGLAIGAIFAATDS
VCTLOVLSODETPLLISLVFGGVNDATSVLNFNAIQSFDMTSPDKIGLHFTGNFL
YLFSSFLGVIGILLCAYIICKLYFGRHSTDRVALLMMLSYLSYIMAEFLYLSGIL
TVFCGIVMSHTVWNHTVSSRVTRHSFATLSFVAETFFLYVGMADLTDEKRFVK
NSOGLSVAVSSILYLGRAAFVPLSFLSNLAKNNSDKISFRQIIILWAGLMR
GAVSTALAYNFTTSGHTSLHENAIMITSVTVLFTSVFGLMTKPLINLLPPHKQ
MPGSHSMTTSEPSPKHFTVPLLDNQDSESDMITGPEVARPTALRMLLRTPTTHVH
RYWRKDDSFNRVFGGRGFVPFVAGSPVEQSPR"
polyA_signal
6138..6143
polyA_signal
6173..6178
polyA_signal
6184..6189
BASE COUNT 2268 a 1241 c 1325 g 2711 t
ORIGIN
```

[illegible]

95 95
3063 GTTCTTAGAGTAGTATTGATTCCTCAACTGGAATTTGTTGGTAATC 3112
95 95
3113 TCGTGGTTCAGTGTGAATTCGCCAAATAAGTGTTTTTTTTTTTTTT 3162
96 PheGlnValLysLysGlnPheP 104
3163 TTTAAATGATATACCTATACAGGTTTCAGTGAANAAGACAGTTT 3212
104 heArgAsnPheValThrIleMetLeuPheGlyAlaValGlyThrIle 120
|||
3213 TCGTGAACCTTCATGACAAATATGCTGTTGGAGCTATTGGCACACTATT 3262
121 SerCysThrIleIleSer Leu 127
3263 AGCTGTTCTATATATCATTTGGTACTAATCAAAACCTCGACATCTTAT 3312
127 127
3313 TACCAGCTAATGCTCTTTGTCCTCAAAATAATTTATGCTGGCATGTG 3362
128 GlyValThrClnPh 132
3363 TATTATCAACTCATATGCGGTATCTATATGATTAGGTGCGGTCAAAAT 3412
132 ePheLysLysLeuAspIleGlyThrPheAspLeuGlyAspTyr 146
: |||
3413 TTTCAAGCACTTAGACATTCACCTTCTGGATTTTGGAGATTTATTAGSTA 3462
146 146
3463 TCATGTATATTCGCTTCTTTACATACACACAGATATCTTCATGCTCT 3512
147 Leu AlaIle 149
3513 TGATTCACAACTTTTTTTTCTCTCTTTTATTCCAAATTCAGCAATT 3562
150 GlyAlaIlePheAlaAlaThrAspSerValCysThrLeuGlnValLeuAs 166
|||
3563 GGTGCGATATTGTCGAACCGATTCTGTTGGACATTCGAGGTACTGAA 3612
166 n 166
3613 TATTATGATTTATGAACGTAGCTATTCAATTCGGTGAAGGCACATTATA 3662
167 GlnAsp 168
3663 TTATAGCTAAATAAATTTTATTTATGCGACAGGTGCTCAGTCAGGAT 3712
169 GluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValValAsnAs 185
|||
3713 GAGAGCCCTTACTTTACAGTCTCGTGTGGAGAGGGGTCTGCTCAATGA 3762
185 pAlaThrSerValValPheAsnAlaIleGlnSerPheAspLeuThrH 202
|||
3763 TGCTACATCTGTGTCCTCTTTTATGCTATTCAAGTTTTTGACATGACTA 3812
202 IsLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTyrLeu 218
: |||
3813 GTTTTGATCCAAAATTTGGGCTTCATTTTCAATGGAAACTTCTTGTAATTA 3862
219 PheLeuLeuSerThrLeuLeuGly 226
|||
3863 TTTCTCTGAGCACTTTTGGGCGTGGAGTAAGTCTCTCTCCGACTA 3912
226 226
3913 GCTATCTCTACCCCGGTGTTGTGCTCTATCCCGCTGTTTCTGTGTC 3962
227 AlaAla 228

3963 TTTACATTTGCTCTTTTCAGCTGCAATTTGACTTCTTATTTCATGCTCTCGCA 4012
: |||
229 ThrGlyLeuIleSerAlaTyrValIleLysLysLeuTyrPheGly 243
|||
4013 GATTGGACTGCTTTGCTGCTATATTATCAAAAGCTATATCTTTGGCAGGT 4062
243 243
4063 AAATCAGTGTATAAAATCTCTATGAGAGTCGGTTTCCTTCCAACTCTTT 4112
243 243
4113 GGTGTTTTTAACCTGTCGTTTTGACCTCTTAGTAAATTTCAATTTATTA 4162
244 Arg HisSerThrAspArgGluValAlaLeuMetMe 255
|||
4163 TATGTGATGGATGACAGGCACCTCAACCGATCGTGGAGTTGCCCTTATGAT 4212
255 tLeuMetAlaTyrLeuSerTyrMetLeuAlaGlu 266
|||
4213 GCTCATGCTCTTACTTGTCTATATAATGCGGAGGTATGATCTAGTTAA 4262
266 266
4263 GTTGTGACGCTAATAATGTTGCTGCTAGTGTGAACTACTTGTCTTCTGTT 4312
267 Le 267
4313 TAAACAAATTTCTGCTGTTGCAATTTAGTAAACTTTTTCGCGTTTTGCACT 4362
267 u.PheAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMet 283
: |||
4363 TATTCTATCTAAGCGGCATCTTACTGTGATTTCTTGTGGAAATTTGTCATG 4412
284 SerHisTyrThrTrpHisAsnValThrGluSerSerArgIleThrThr ... 299
|||
4413 TCTCATTTATACCTGGCACAATGTTACCGAGAGCTCAAGGTCACACTACTAG 4462
299 299
4463 GTATGCTCTACAAACCTCTCTTCGAGCTCTTTGCAAGTTCTCGTTATTT 4512
300 Lys 300
4513 AACACAGTTTTGCGAATTTGGTTTAAATCTACTCCGTTTTTGGTTGCGAG 4562
301 HisThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTy 317
|||
4563 CATTCCTTTGCAACTCTGTTCATTTGTCGAGAGACATTTATCTCTCTCTA 4612
317 rValGlyMetAspAlaLeuAspIleAspLysTrpArgSerValSerAsp. 333
|||
4613 TGTGATGATGGATCGCTTGGATATCAGAAATGGAATTTGTGAAAAATA 4662
333 333
4663 GGTATGTTGAATACGGGTTACACCTATAGTCTTGTGATGCTTGAATGAG 4712
333 333
4713 CAAGAAATAATTTGTTGCTCAATCAGAAGTATTCACATATTCCTCCCTCC 4762
334 ThrPro 335
4763 AAAATAAGTTATAAACTATTGTTGGTCAACGCCAGTTAACTTCAAGAAATTT 4812
335 335
4813 AGTTCTTGTTAAGAGAGAGAAAAACAAAAACATATGAGTTTATAGGT 4862
335 335

4863 CATGAGTTAGACTAACTAAATTTGATTGGATCGAGTCCTTTTAGTGGTTTA 4912
335 335
4913 TTATAAGTTATAACAATTTTAAAGATAAAACAATAAACTATGTTCCAG 4962
336GlyThr 337
4963 GAACCTTATAAATCTCTCTGTTATTATTATTCTTTCAGTCAGGGACTA 5012
338 SerIleAlaValSerSerIleLeuMetGlyLeuValMetValGlyArgAl 354
5063 TCGGTCGTTATTCCTCTGTCGTTTATTCCTCACTTAGCAAGAACT 5112
5013 TCAGTTGACGTAGCTCAATATTTGTTAGCCCTAATCTTAGTACGACAGC 5062
354 aAlaPheValPheProLeuSerPheLeuSerAsnLeuAlaLysLysAsnG 371
5063 TCGGTCGTTATTCCTCTGTCGTTTATTCCTCACTTAGCAAGAACT 5112
371 InSerGluLysIleAsnPheAsnMetGlnVal..Val..... 382
5113 CTTGCGACAAGATATCTTTAGGCAACAAGTAAGGAGTCTCTACAAAT 5162
382 382
5163 TCTCATTTTATATTATCTTTACGAGTAATATACATCTCTTACTCTCTGTC 5212
383IleTrpTrpSerGlyLeuMetArgGlyAlaValSerMe 395
5213 TCTTAGATATAATTTGGTGGCTGCCTAATGAGAGCGCGTCTCAAT 5262
395 tAlaLeuAlaValAsn..Lys..... 401
5263 AGCACTTGGTATAATTAAGGTAAGTATAATATAGTCACACAAATGTTA 5312
401 401
5313 CAATGTCACCTTAGAATTGAAGAAGCGGTGGTGGTCTTCTTCATTTCC 5362
402PheT 403
5363 CGATTTCTTGTTGATCTGTTAGATTAGATGTTTATGACGCGGTGCAGTTTA 5412
403 hrArgAlaGlyHisThrAspValArgGlyAsnAlaIleMetIleThrSer 419
5413 CAACCTCGGGCATAGCTATTCACGACGACGCAATATATGATCAAGT 5462
420 ThrIleThrValCysLeuPheSerThrVal..... 429
5463 ACTGTTACGGTGTCTGTTTCAGCAGAGTTGTAAGGACCTACATACAT 5512
429 429
5513 TTTTCGAGGCTTTTCTTTGATTCTCGCCAAATGAATCTTACGTTGG 5562
430ValPheGlyMetLeuThr 435
5563 TTGTTGGTGTGTTGTTGGTGTGTTTATGATATTTCGGTGTGATGACG 5612
436 LysProLeuIleSerThrLeuLeuPro..... 444
5613 AAGCCTCTGATAACCTTCTGCTACCCCGCACCAAGCAGATGCCAAGCGG 5662
445 HisGlnAsnAlaThrThrSerMetLeuSerAspAsnThrProLys 461
5663 TCATTCGTCATGACACA.....TCCGAACCCAGTAGTCCGAAGC 5703
461 erIleHisIleProLeuLeuAspGln.....AspSerPheIleGluPro 475
5704 ACTTCACGGTGCCACTCCTCGACACCAACCACTGACTCAGAAACGATATG 5753
476 SerGlyAsnHisAsnValProArgProAspSerIleArgGlyPheLeuTh 492
5754 ATAACCGGACCTGAGGTGCTCGACCAACCTGCCTTGGCCATGCTGCTAAG 5803

492 rArgProThrArgThrValHisThrTyrTrpArgGlnPheAspSerP 509
5804 GAGCCCAACCCACACGCTGACCGCTACTGCGTAAGTTTGATGATCGT 5853
509 heMetArgProValPheGlyGlyArgGlyPheValProPheValProGly 525
5854 TTATGCGTCCGTTGTTGGCGGGGATTCGTTCCGTTTTCGCGGGC 5903
526 SerProThrGluArgAsnPro 532
5904 TCACCAAGTTGACAGAGCCCT 5924
seq_name: gb_p11:AB015479
seq_documentation_block: 80675 bp DNA PLN 09-AUG-2000
LOCUS AB015479 Arabidopsis thaliana genomic DNA, chromosome 5, pl clone:MTE17.
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, pl clone:MTE17.
ACCESSION AB015479
VERSION AB015479.1 GI:3241927
KEYWORDS Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1
SOURCE clone:MTE17.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (sites)
AUTHORS Nakamura,Y., Sato,S., Asamizu,E., Kaneko,T., Kotani,H., Miyajima,N.
and Tabata,S.
TITLE Structural analysis of Arabidopsis thaliana chromosome 5. VII.
Sequence features of the regions of 1,013,767 bp covered by sixteen
physically assigned P1 and TAC clones
JOURNAL DNA Res. 5 (5), 297-308 (1998)
MEDLINE 99087489
REFERENCE 2 (bases 1 to 80675)
AUTHORS Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-1998) to the DDBJ/EMBL/GenBank databases.
Yasukazu Nakamura, Kazusa DNA Research Institute, Department of
Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:ynakamu@kazusa.or.jp, Tel:81-438-52-3935,
Fax:81-438-52-3934)
COMMENT Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MTE17
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/Grail-1.3/),
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volker Brendel, Stanford University,
http://gremlini.zool.iastate.edu/cgi-bin/sp.cgi).
Genes encoding tRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
This sequence may not be the entire insert of this clone. It may be
shorter because we remove overlaps between neighboring submissions.
The 5' clone is MCOL5 and the 3' clone is MWC10.
FEATURES
source
1. 80675
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="5"
/clone="MTE17"
/clone_lib="Mitsui P1"
complement(join(636..713,796..964,1048..1151,1258..1347,
1437..1505,1645..1778,1867..1957,2059..2139,2265..2390,

Ratio: 2.529 Gaps: 24
Percent Similarity: 43.719 Percent Identity: 30.762

alignment_block:

US-09-271-584a-2 x AB015479/rev ..

Align seg 1/1 to reverse of: AB015479 from: 1 to: 80675

```
15 SerAspHisAlaSerValValAlaLeuAsnLeuPheValAlaLeuLeuCy 31
   ::::::::::: ::::::::::: ::::::::::: :::::::::::
48342 GCTGAGCATCTCAGGTGATACCAATCTCAGTGTTCATCGCCATCTCTG 48293
31 salacysileValLeuGlyHisLeuLeuGluGluAsnArgTrpMetAsnG 48
   ::::::::::: ::::::::::: ::::::::::: :::::::::::
48292 TCTATGTTTAGTTATCGCCACCTGCTTGAAGAGAATCGATGGGTTAATG 48243
48 luserIleThrAlaLeuLeu..... 54
48242 AATCTATTACCGCCATTTTAGTAGTAATTTCTTCAAAATTTCTTCTTTT 48193
54 ..... 54
48192 TTTTGTTTAGCTAAAGTTTTCAGCTTTTCGCTCTTAAAGAACAAATTT 48143
55 ..... Ile..GlyLeuGlyThrGlyVa 61
   ::::::::::: ::::::::::: ::::::::::: :::::::::::
48142 ATCAGTATTTTTCATGTTTGTTCAGTAGGAGGAGCAGCATCAGGAAC, 48093
61 lThrIleLeuLeuIleSerLysGlyLysSerSerHisLeuLeuValPheS 78
   ::::::::::: ::::::::::: ::::::::::: :::::::::::
48092 AGTACTCTTACTATTAGTAAAGGAAAAAGTTTCACATATTTTGTGTTTG 48043
78 erGluAspLeuPhePheIleThrLeuLeuProProIleIlePheAsnAla 94
   ::::::::::: ::::::::::: ::::::::::: :::::::::::
48042 ATGAAGAAGCTCTCTTCATTTTACCTTCCTCCCAATATCTTCAATGCT 47993
95 .Gly..... 95
   |||
47992 GGGTAAGCTTACTATGTTTGGTTACTATGTTGGTTAATGGGTTAGA 47943
95 ..... 95
47942 TTACACTATGATGGTGATTGACTGAAGATTTTGGTTTGTGGCAGG 47893
96 PheGlnValLysLysGlnPhePheArgAsnPheValIleMetLe 112
   ::::::::::: ::::::::::: ::::::::::: :::::::::::
47892 TTCCAAAGTTAAGAAAAGAGTTTTCACAACTTTTAAACCATCATGTC 47843
112 uPheGlyAlaValGlyThrIleIleSerCysThrIleIleSerLeuGlyV 129
   ::::::::::: ::::::::::: ::::::::::: :::::::::::
47842 CTTTGGTGTGATGGAGTTTTCATCTCCACTGTCAATTATCTCGTTGTA 47793
129 alThrGlnPhePheLysLysLeuAspIleGlyThrPheAspLeuGly...A 145
   ::::::::::: ::::::::::: ::::::::::: :::::::::::
47792 AAAAAATTTTCAATTAATCTTGAATGACTTGACTATATGATTTAAAGTTG 47743
145 spTyrLeu..... 147
   |||
47742 ATTATCTCGAGTTGTACCACTCTATTTTGTCTTTTATATATGCA 47693
147 ..... 147
47692 ACCAGGACTTGTGGCTGTGTTCCCAAGTTGGGATTTAAGGGGTTGAGTG 47643
147 ..... 147
47642 CTAGAGACTATCTTGTGAAGATTTTTCAGTGAATGTTCTTTTCAATT 47593
147 ..... 147
47592 TCTTGTGAGATGTTATTACTTATCAGTTTTCGAGAGTGAATGCTTTCT 47543
148 .....AlaIleGlyAlaIlePheAlaAlaThrAspSerValCy 160
```

```
|||||:::|||||:::|||||:::|||||:::|||||:::||||| 47542
TGTTTTGATTTCAGCCATAGGAACGATTTTCTCATCAACTGATACTGTTG 47493
160 sThrLeuGlnVal..... 164
|||||:::|||||
47492 CACTCTACAGGTTCTTGATTTAGAAGATTCTTTGTCTTACTACTTCTCT 47443
165 .....LeuAsn 166
|||||
47442 ATTAAAGTTTTTCTTACATTCATCTTAAACTTTTACAGATTCTCCAT 47393
167 GlnAspGluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValVa 183
|||||:::|||||:::|||||:::|||||:::|||||
47392 CAAGATGAACACCATTTGCTATACAGCTTAGTCTTTGGAGAAGGAGTGT 47343
183 lAsnAspAlaThrSerValValPheAsnAlaIleGlnSerPheAspL 200
|||||:::|||||:::|||||:::|||||:::|||||
47342 GAATGATGAACCTCAGTTGTACTGTTCACCGCGGTGCAAAAGATTCAAT 47293
200 euThrHisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeu 216
:::|||||:::|||||:::|||||:::|||||
47292 TTGAAGCCTAACCGGTGGACGCGCTGCAAGTATTTTGGAACTTTTG 47243
217 TyrLeuPheLeuLeuSerThrLeuLeuGly..... 226
|||||
47242 TAGCTCTCTCTCAACAAGCACACTTCTCGGAATTTGGTAAAGCTATTGAGT 47193
226 ..... 226
47192 TGGCATAGTTTTTCTCTGCTTTGAAGCATATCATGAGTAATCTTAGAAC 47143
227 .....AlaAla..ThrGlyLeuIleSerAlaTyrValIleLysLysL 240
   ::::::::::: ::::::::::: ::::::::::: :::::::::::
47142 TGTGAATGCTCTGCAGGTGGGCTTAATACATCTTTTGTCTTAAACCT 47093
240 euTyrPheGly..... 243
   |||
47092 TGTATTTTGGAGGTAAAGTACACCAAACTTGCAAAATTTTTCATCTCAA 47043
243 ..... 243
47042 GCTTAAGTTTTTTTTTTTACTGATAGATGCTGCTTTATATTGATCTTG 46993
244 .....ArgHisSerThrAspArgGluValAlaLeuMetM 255
   |||
46992 TTATATGTTGCTGCACAGACATTTCTACTACACGCAACTGCCATCATGG 46943
255 etLeuMetAlaTyrLeuSerTyrMetLeuAlaGlu..... 266
   ::::::::::: ::::::::::: ::::::::::: :::::::::::
46942 TTCTAATGGCTTACCTTTTCATATATGTTGGTGAGGTAATACACTGGTCA 46893
266 ..... 266
46892 TGAACGCTACTTTTTTAGTTTTCATTTTGAATTTGAAGTGTGCTGCTCTCT 46843
267 .....LeuPheAspLeuSerGlyI 273
   |||
46842 GTTTTAACTCAATATAGTTGCTCTCTTCAGCTCTCTCTCAATAAGTGAA 46793
273 leLeuThrValPhePheCysGlyIleValMetSerHisTyrThrTrpHis 289
   |||
46792 TCTTACTGTTTTCTCTGTTGGTGTGTTTAAATGTCGCATTTATGCATCATAT 46743
290 AsnValThrGluSerSerArgIleThrThr..... 299
   |||
46742 AACGTGACAGAGAGCTCAAGAAATCACTTCCAGGCAATGTATCAGACAATT 46693
299 ..... 299
46692 GTTTATTGCGTGAAGTCTCAATTTTCAGGAATTCCTTTTGTCTGACTGTG 46643
300 .....LysHisThrPheAlaThrLeuSerPheLeuAlaGlu 311
   ::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

```
45642 CATCTATATTTTCAGCGATGATTGCAATGTTGCTCTTTTATTCGGGAG 46593
312 ThrPheIlePheLeuTyrValGlyMetAspAlaLeuAspIleAspLysTr 328
|||||
46592 ACATTCATATTTCTGTATGTTGGACAGATGCTCTGTGATTTCACAAAGTG 46543
328 pArgSerValSer..... 332
|||||
46542 GAAGACAAGCAGCTTAAGGTATAAAGTCGTATCCATTTATCAGCTGGTAG 46493
332 ..... 332
46492 AGCGTTTAGAGTGTGTTGTCAAAACAATACACAGCATTTTGTATATGAACC 46443
333 ..... Asp.ThrProGlyThrS 338
|||||
46442 TCGTGTATTTCTTAATGCATACCTAGATCAAGATTATCCCAAGACAT 46393
338 er..... 338
46392 CCTCTATGCTCTTCTCTCTCTTGACACGAAGTTAATCGCAATTAACTGT 46343
338 ..... IleAlav 341
46292 AGACITGTAAACATTAAATCATCTCTTCGACGTTTGGGGGTACTCTGGGTG 46243
341 alSerSerIleLeuMetGlyLeuValMetValGlyArgAlaAlaPheVal 357
|||||
46242 TCTCCGGTGTATACCCGATTTAGTATTGCTTGGACGAGCAGCATTTGTC 46193
358 PheProLeuSerPheLeuSerAsnLeuAlaLysLysAsn.....GlnSe 372
|||||
46192 TTTTCACCTCTCGGCTTAAACAATTTTCATGAACAGGCACACAGTGAAGAAA 46143
372 rGlutLysIleAsnPhe.AsnMetGln..... 380
46142 CGAGTCTATACATTTAAGCATCAGGTTTACTCTTGTGTTTACAGTTA 46093
380 ..... 380
46092 CATTATCAGAAGCAGCTAAATAAATAAATAAATTAATGTCGCCAGATGGGA 46043
381 .....ValValIleT 384
46042 AAGTTTCTGTTATATAAATCTCTCTCTTGTATACATTTATAGTGATCATTT 45993
384 rpTrpSerGlyLeuMetArgGlyAlaValSerMetAlaLeuAlaTyrAsn 400
|||||
45992 GGTGGGCAGGCTAATGCGAGTGCTGCTCAATTGCTGCGCTTCAAG 45943
400 ..... 400
45942 CAGGTTTTCTCTTTGACATTCATATTAATCTCTAAATCTCTATTGTATTA 45893
400 ..... 400
45892 CTTTACATGAATTAGTAACATTTCCATTTATATCTCATTTTGGTATATC 45843
401 .....LysPheThrArgAlaGlyHisThrAspValArgGlyAsn 413
|||||
45842 TTTTCATGCGCAAGTTCATACACTCTCCGGTGTATACATTTGGATCCTGTGAAT 45793
414 AlalleMetIleThrSerThrIleThrValCysLeuPheSerThrVal... 429
|||||
45792 GCTGCCATGGTCACCAACACCATCATGTTGTTCTCTTCTTACTACACTGGT 45743
429 ..... 429
45742 AAACAATACTACCTTAGTATATATGGCCCTACTCGTCTGCTAGTATTTCT 45693
```

```
430 .....ValPhe 431
45692 GTGTTAAGATCAAAAGAAATGACTTTTGTTCATGTAACAAACAGGTCTTT 45643
432 GlyMetLeuThrLysProLeuIleSerTyrLeuLeuPro..... 444
|||||
45642 GGTTCCTCACAACACCATGTCGAATATCTCTTCTCAAGATGCAAG 45593
445 HisGlnAsnAlaThrSerMetLeuSerAspAsnThrProLys. 460
|||||
45592 TCACAACACCGGAATAGAGTAACACGACTGACCCAGGTTCTCCGAAG 45543
461 SerIleHisIleProLeuLeuAspGlnAspSerPheIleGluProSer 476
|||||
45542 AAGATGCGACACTTCTCTCTCTT.....TCCTTTGACGAGTCTGCT 45502
477 GlyAsnHisAsnValProArgProAspSerIleArgGlyPheLeuThrAr 493
|||||
45501 TCCACCAACTTCATAGACTAAGATAGTATTTCCTCTCTGATGGAACA 45452
493 gProThrArgThrValHisTyrTyrArgGlnPheAspAspSerPheM 510
|||||
45451 ACCTGTGTACACCATCCACCGCTACTGGAGAAAGTTTGACGACACATACA 45402
510 etArgProValPheGlyArgGlyPheValProPheValProGlySer 526
|||||
45401 TGAGGCTATCTTTCGGTGA..... 45382
527 ProThrGluArgAsnProProAsp 534
|||||
45381 CCTCGTCGAGAAAACCAACAGAA 45358
seq_name: gb_in2:AF142676
seq_documentation_block: 2405 bp mRNA INV 27-MAY-1999
LOCUS AF142676 Drosophila melanogaster sodium-hydrogen exchanger NHE1 (NHE1) mRNA,
DEFINITION complete cds.
ACCESSION AF142676
VERSION AF142676.1 GI:4894990
KEYWORDS fruit fly,
SOURCE Drosophila melanogaster
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 2405)
AUTHORS Dow,J.A.T.
TITLE An insect member of the Na+/H+ exchanger family
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2405)
AUTHORS Dow,J.A.T.
TITLE Direct Submission
JOURNAL Submitted (14-APR-1999) IBLS-Molecular Genetics, University of
Glasgow, Anderson College, Glasgow, Scotland G11 6NU, UK
FEATURES
source
1..2405
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
1..2405
/gene="NHE1"
32..1717
/gene="NHE1"
/codon_start=1
/product="sodium-hydrogen exchanger NHE1"
/protein_id="AAD32689.1"
/db_xref="GI:4894991"
/translation="MRWVAYSAALLLVHAGPESIGQEVPOSKYSSNTTDDNSS
SIHTVDVFNPSIGNVTPTSIASGNASTTKRGNASTLVDPDPLIDSHAVEQEHNSSL
SLEFVICVIMLGILLIHSMLQTFQYLPESIVVFLGAFGLSLNVMGQNGMKREE
VFSPMGFLLPPTIFESGYNLHKGNFQNGISILVFAIFGTTISALVTGAGIYLLG
LGEVAFRLSFSESAFGLISAVDPVATVAFIHALDVPDILNMLVFESILNDAISIV
```



```

mRNA, partial cds.
ACCESSION AF279671
VERSION AF279671.1 GI:9652169
KEYWORDS common ice plant.
SOURCE Mesembryanthemum crystallinum
ORGANISM Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales;
Caryophyllales; Aizoaceae; Mesembryanthemum.
REFERENCE 1 (bases 1 to 421)
AUTHORS Chauhan,S., Forsthoefel,N. and Bohnert,H.J.
TITLE Na+/myo-inositol symporters and Na+/H+-antiporters in plant sodium
transport
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 421)
AUTHORS Chauhan,S., Forsthoefel,N. and Bohnert,H.J.
TITLE Direct Submission
JOURNAL Submitted (19-JUN-2000) Biochemistry, University of Arizona, 1041
E. Lowell St, Tucson, AZ 85721, USA
FEATURES
source
location/Qualifiers
1..421
/organism="Mesembryanthemum crystallinum"
/db_xref="caxon:3544"
<1..>421
/gene="NHX2"
<1..>421
/gene="NHX2"
/codon_start=1
/product="putative Na+/H+-antiporter"
/protein_id="AAF91424.1"
/db_xref="GI:9652170"
/translation="KKQFFRNFTIIVFGALSTLVSTIISLGAMEFFKKLDIGSLEL
GDYLAIGAFNATDSVCTLOVLHODEPPLYSIVFGGVNDATSVVLFNATQNFDTL
HIDRIAFKFNAGNLYLFTSTLGLATGLLSAYLIKK"
BASE COUNT 110 a 85 c 88 g 138 t
ORIGIN
alignment_scores:
Quality: 586.00 Length: 140
Ratio: 4.406 Gaps: 0
Percent Similarity: 95.000 Percent Identity: 80.000
alignment_block:
US-09-271-584A-2 x AF279671 ..
Align seg 1/1 to: AF279671 from: 1 to: 421
100 LysLysGlnPhePheArgAsnValThrIleMetLeuPheGlyAlaVal 116
|||||
1 AAGAAAGCAATTTTCCGGAATTCATTACAATCATAGTATTGGAGCCCT 50
|||||
116 lGlyThrIleSerCysThrIleIleSerLeuGlyValThrGlnPheP 133
|||||
51 GGGTACATTGGTGTCTATTTACCATCATATACACTAGGAGCTAGGAGTCT 100
|||||
133 heLysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrIleAlaIle 149
|||||
101 TCAGAAAGTTGGACATTGGTGTCTAGAGTTAGGAGCTATCTTGCATTT 150
|||||
150 GlyAlaIlePheAlaIleThrAspSerValCysThrLeuGlnValLeuAs 166
|||||
151 GGTGCATATTCGCTGCACACCGACTCTGTTGCACGCTGCAGCTGCTCA 200
|||||
166 nGlnAspGluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValV 183
|||||
201 TCAGGATGAGACCCCTCTACTATACAGTCTGGTGTGGGGAAGGTGTAG 250
|||||
183 alAsnAspAlaThrSerValValValPheAsnAlaIleGlnSerPheAsp 199
|||||
251 TCAATGATGACATCATAGTGGTCTTTTATGCCATCAGACTTTTGAC 300
|||||
200 LeuThrHisLeuAsnHisGluAlaAlaPheHisLeuLeuGlyAsnPheLe 216
|||||
|||||
301 CTCACATATCGATCACGAATTCCTTCAATTTGCTGGCACTTTT 350
|||||
216 uTyrLeuPheLeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIleS 233
|||||
351 ATATTTATTTTACACGACGCTGCTGGGCAATTTACTGGATTGCTCA 400
|||||
233 erAlaTyrValIleLeuLysLys 239
|||||
401 GTGCATACATTATCAAAAG 420
|||||
seq_name: gb_pr5:D87743
seq_documentation_block:
LOCUS D87743 4408 bp mRNA PRI 10-JUL-1997
DEFINITION Human mRNA for KIAA0267 gene, partial cds.
ACCESSION D87743
VERSION D87743.1 GI:1665826
KEYWORDS KIAA0267.
SOURCE Homo sapiens male bone marrow Myeloblast cell_line:KG-1 cDNA to
mRNA, clone_lib:library in pBluescript II SK clone:HA7045.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4408)
AUTHORS Nomura,N.
TITLE Direct Submission
JOURNAL Submitted (05-SEP-1996) to the DDBJ/EMBL/GenBank databases. Nobuo
Nomura, Kazusa DNA Research Institute, Gene Structure 1; 1532-3
Yana, Kizarazu, Chiba 292, Japan (E-mail:cdna@kazusa.or.jp,
URL:http://www.kazusa.or.jp, Tel:0438-52-3930, Fax:0438-52-3931)
2 (sites)
REFERENCE Nagase,T., Seki,N., Ishikawa,K., Ohira,M., Kawarabayashi,Y.,
AUTHORS Ohara,O., Tanaka,A., Kotani,H., Miyajima,N. and Nomura,N.
TITLE Prediction of the coding sequences of unidentified human genes. VI.
The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
analysis of cDNA clones from Human cell line KG-1 and brain
analysis of cDNA clones from Human cell line KG-1 and brain
Unpublished (1996)
3 (sites)
REFERENCE Nagase,T., Seki,N., Ishikawa,K., Ohira,M., Kawarabayashi,Y.,
AUTHORS Ohara,O., Tanaka,A., Kotani,H., Miyajima,N. and Nomura,N.
TITLE Prediction of the coding sequences of unidentified human genes. VI.
The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
analysis of cDNA clones from cell line KG-1 and brain
DNA Res. 3 (5), 321-329 (1996)
97191544
FEATURES
Location/Qualifiers
1..4408
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="KG-1"
/cell_type="Myeloblast"
/clone_lib="HA7045"
/sex="male"
/tissue_type="bone marrow"
1..2001
/gene="KIAA0267"
<1..2001
/note="Similar to Human Na+/H+ exchanger 2 (A57644)"
/codon_start=1
/protein_id="BAA13449.1"
/db_xref="GI:1665827"
/translation="RGWRAPLRGVSSPAPRLMRPLMLLLAVGVDFWAGASDGG
GEARMDDEIVSEKQAEESHRQDSANLLIFLLTLTLTLWLFKRRARFLHETGLA
MTYGLLVGLVRYGHVPSDVNNVTLSCEVSSPTLLVTFDVEFNILLPTIIFYA
GYSKRRHFRFNLGSLAYAPLGTATSCFVIGSIMYCVTKMKTGVLADGFTDCL
LFGATVSATDPVTVAIFHELOVDVLYALLFESVLDVAIVLVSSIVAYOPAGDN
SHTFDVTAMFKSIGIFIGISFGSFGAMGATGVVTVLTFTKTLREQLLETGFFLMS
WSTELLAEAWGFTGVAVLFCGITQAHYTYNNLSTESQHETKQLFELLNENFLFS
YNGLFTFTQNHVFNPTVVGAFVAILFGRANILYPLSLNLLNLRGRSKIGSNFQHMNM
FAGLRGAMAFALINDTATYARQMMFSTLLIVFTVWVFGGTTAMLSCILIRVGD
SDQHLGVPENERERTTAEASAWLFRRMYNFDHNYLKPILLTHSGPPLTTLTLPACCGPIA

```

us-09-271-584a-2.p2n.rge

```

896  CTTTTGC AATGGGTGCTGCTACTTGGAGTGGTGACACGTTTAGTGC AAG  945
      ||      ::::::::::::::::::::::::::::::::::::|||      |||
239  ....LysLeuTyrPheGlyArgHisSerThrAspArgGluValAlaLeu  253
      |||||      |||      |||      |||      |||      |||
946  TTCACCAAAATTA.....CGGGAGTTCACGTTCTTGGAGACAGCGCT  986
      |||||
253  uMetMetLeuMetaLalTyrLeuSerTyrMetLeuAlaGluLeuPheAspL  270
      |||      |||      |||      |||      |||      |||
987  GTTCTCTTCATGCTCCGGAGTACCTTCCTTCCTGGCTGAAGCATGGGCT  1036
      |||      |||      |||      |||      |||      |||
270  euSerGlyIleLeuThrValPhePheCysGlyIleValMetSerHisTyr  286
      |||      |||      |||      |||      |||      |||
297  TCACACGCTGACATTCGCTACTTCTTTGGCATCATCACAAAGCACATAT  1086
      |||      |||      |||      |||      |||      |||

```

```

253 uMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPheAspL 270
      : : : : : : : : : : : : : : : : : : : : : : : : : :
987 GTTCTTCTTGTCTGGAGTACCTTCTCTTGGGTGAAGCATGGGGCT 1036
      : : : : : : : : : : : : : : : : : : : : : : : : : :
270 euSerGlyIleLeuThrValPhePheCysGlyIleValMetSerHisTyr 286
      : : : : : : : : : : : : : : : : : : : : : : : : : :
987 TCACACCTCTGACCTGACCTATCTTTTGTGGCATTCACACAGCACATAT 1086
      : : : : : : : : : : : : : : : : : : : : : : : : : :

```

270 euSerGlyIleLeuThrValPhePheCysGlyIleValMetSerHisLyr 286
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
037 mCACACCTCTNCTTGCACATATCTCTTTTGTCATCACACAAGCAATTAT 1086

COC 970-861-111

[illegible]

1642ACCCACAGCGGGCCTCCG 1659

seq_name: gb_pr3:AF030409

seq_documentation_block:			
LOCUS	AF030409	4452 bp	mRNA
DEFINITION	Homo sapiens sodium-hydrogen exchanger 6 (NHE-6) mRNA, nuclear gene encoding mitochondrial protein, complete cds.		
			PRI 09-MAR-1998

AF030409.1 GI:2944232

KEYWORDS . human.
SOURCE

ORGANISM HOMO SAPIENS
Eukaryota: Metazoa

REFERENCE
1 (bases 1 to 1152)
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS
Numata, M., Petrecca, K., Lake, N. and Orłowski, J.

JOURNAL
J. Biol. Chem. 273 (12), 6951-6959 (1998)

HEADLINE
38173903
REFERENCE 2 (bases 1 to 4452)

AUTHORS	Numata, M. and Orlowski, J.
TITLE	Direct Submission

JOURNAL Submitted (18-OCT-1997) Physiology, McGill University, 3655

COMMENT This mRNA sequence extends the sequence in

FEATURES	Location/Qualifiers
Accession Number D87743.	

source
1. .4452
/organism="Homo sapiens"

```
/db_xref="taxon:9606"
```

```

/ gene="NHE-6"

```

CD3 30. :2043 /gene="NHE-6"

```

/function="ion transporter
/codon_start=1

```

```

/product="sodium

```

```
/db_xref="GI:2944233"
```

GGGGEARAMDEETVSEKOAEEF
/CLAUSIOLII= MARRGW

GLAM¹YGLLVGLVLR²YGIHV³F⁴
FYACVSIKPBHEEDNYCS⁵TI⁶NA⁷

DCLLFGAIVSATDPVTVLAIF

LMSWSTFLAEAWGETGVAV

IFSYMGLTFFTFQNHVFNPTE
MMMFAGLRGAMAFALAIRDTA

GVDS DQ E H L G V P E N E R R T T K A
D I A R C I T S D P A V E N O T O Y K D D

SEDALDRELAFGDHELVI RGT

ORIGIN

alignment_scores:	
Quality:	591 00
Count:	591 00

Ratio:	Gaps:
1.699	20

0001:00 : Parameter success? : Parameter success? : Parameter success?

US-09-271-584A-2

Align seq 1/1 to: NC030400

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

|| ::||| ||||| ::|||:: |||

253 CAGCCTATGTCATCCGCGTGCACCCCTACCATTCTCACAAATCTGGCT 309

39 uLeuGluGluAsn.....ArgTrpMetAsnGluSerIleThrAlaLeuL 54

305 CTTCAAGCACCGCGCGCGCGCTTCTTGCACGAAACCGCGCTGGCTATGA 354

54 euIleGlyLeuGlyThrGlyValrHrIleLeuLeu.....IleSerLysGlyLysse 71
 :
 355 TTTATGGTCTTTTGTGGGCCTTGTCGTCGGTATGCCATTGATGTTCCG 404
 66IleSerLysGlyLysse 71
 :
 405 AGTGATGTAATAATATGTACCCCTGAGCTGGAAGTGCAGTCAAGTCCAAC 454
 71 rSerHisLeuLeuValPheSerGluAspLeuPhePheIleTyLeuLeuLeu 88
 :
 455 TACCCTACTGCTTACTTTGTATCCAGAAGTATTTTTCAACATATTACTTC 504
 88 roProIlePheAsnAlaGlyPheGlnValLysLysGlnPhePhe 104
 :
 505 CTCCTATCATATTATCGCAGGTATATAGCTGAAAGGAGACATTTTTT 554
 105 ArgAsnPheValThrIleMetLeuPheGlyAlaValGlyThrIleIleSe 121
 :
 555 CGAATCTGGGTCATCCCTAGCATAGCTTTCTTGGAACACAACAATTTC 604
 121 rCySerThrIleIleSer.....LeuGlyValThrGlnPhePheLysL 135
 :
 605 TTGTTTCGTTATTTGGTCAATATGCTGTGTAAACGCTGATGAAG 654
 135 ysLeuAsp.....IleGlyThrPheAspLeuGlyAspTyLeuAlaIle 149
 :
 655 TAACGGGACAACCTGCAGAGATTTTACATTTACAGATTCCTACTGTT 704
 150 GlyAlaIlePheAlaThrAspSerValCySerThrLeuGlnValLeuAs 166
 :
 705 GGTGCCATTGATCAGCAACTGATCCAGTCACTGCTTCTGCTATATCCA 754
 166 nGlnAspGluThrProLeu...LeutySerLeuValPheGlyGluGly 182
 :
 755 CGAGCTTCAAGTTGATGTGAACCTCTATGCACCTCTTTTGGTGAAGTG 804
 182 alValAsnAspAlaThrSerValValValPheAsnAlaIle..... 195
 :
 805 TCCTCAATGATGCTGTGCCATAGTGTCTCCCAATATGATGGCATAC 854
 196GlnSerPheAspLeuThrHisLeuAsnHi 205
 :
 855 CAGCCAGCTGGAGACAACAGTCACACCTTTGATGTACA..... 893
 205 sGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTyLeuPheLeuLeus 222
 :
 894 ...GCGATGTTCAAGTCTATTGGGATCTCTCCCTTGAATCTTCAGTGGAT 939
 222 erThrLeuLeuGlyAlaAlaThrGlyLeuIleSerAlaTyValIleLys 238
 :
 940 CTTTGGCAATGGGTGCTGCTACTGGAGTGTGACAGCTTAGTGACAAAG 989
 239LysLeuTyTPheGlyArgHisSerThrAspArgGluValAlaLe 253
 :
 990 TTCACCAAAATTA.....CGGGAGTTCCAGTTCTGTGGACACAGGCCT 1030
 253 uMetMetLeuMetAlaTyTrLeuSerTyTrMetLeuAlaGluLeuPheAspL 270
 :
 1031 GTTCTTCTGATGTCCTGGAGTACCTTCTCTTGGCTGAAGCATGGGGCT 1080
 270 euSerGlyIleLeuThrValPhePheCysGlyTyleValMetSerHisTy 286
 :
 1081 TCACAGGTGATGTGCACATATTGTTTGGGCATCACACAAGCACCAATTAT 1130
 287 ThrTrpHisAsnValThrGluSerSerArgIleThrThryHisThrPh 303
 :
 1131 ACGTATAATAATTTGCCAGGAGTCTCAGCATAGAACAATAACAGTTGTT 1180
 303 eAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyValGlyM 320
 :
 1181 TGAGCTTCTCAATTTCTTGGCAGAAATTTCACTTCTCTCATATGGGC 1230

320	etAspAlaLeuAspIleAspLysTrpArgSerValSerAspThrProGly	336
1231	TGACACTGTTCCACCTTCCAG.....AACCATGCTTTTAC	1265
337	ThrSerIleAlaValSerSerIleLeuMetGlyLeuValMetValClyAr	353
1266	CCACATTTCTAGTAGGACGATTGTT.....GCTATTTTCTTGGGAAG	1309
353	gAlaAlaPheValPheProLeuSerPheLeuSerAsnAlaLysLysA	370
1310	AGTCGCCAATATTTACCCCTGTCCTCTACTTAATTTGGGTAGAAGAA	1359
370	snGlnSerGluLysIleAsnPheAsnMetGlnValValIleTrpTrpSer	386
1360	GT.....AAGATTGGATCAAAATTTTCAACACATGATGATTGGC	1400
387	GlyLeuMetArgGlyAlaValSerMetAlaLeuAlaTyrAsnLysPheTh	403
1401	GGCCTT...CGTGGTCAATGGCATTTGGCTTGGCCATTTCGA.....	1439
403	rArgAlaGlyHisThrAspValArgGlyAsnAlaIleMetIleThrSerT	420
1440GATACGCCACTTATCGACGGCAATATGATGTCAGGACCA	1479
420	hrIleThrValCysLeuPheSerThrValValPheGlyMetLeuThrLys	436
1480	CGCTTCGATTGCTTTTACCGCTGGGTATTGGT.....	1517
437	ProLeuIleSerTyrLeuLeuProHisGlnAsnAlaThrThrSerMetLe	453
1518GGTGGCACCACCTGCAATGCT	1537
453	uSerAspAspAsnThrProLysSerIleHisIleProLeu.....	467
1538	GTCA.....TGCTTGCATATCACGGTGTGGTTGATT	1569
467	euAspGlnAspSerPheIleGluProSerGlyAsnHisAsnValProArg	483
1570	CAGACCAAGAACACTTGGGTCTCTCAAAATGAAAGAGAACTACCAAA	1619
484	ProAspSerIleArgGlyPheLeuThrArgProThrArgThrValHisTy	500
1620	GCAGAGAGT.....GCTTGGCTTTTCCCG.....	1643
500	rTyTrpArgGlnPheAspAspSerPheMetArgProValPheGlyAla	517
1644	ATGTGTACAACTTGTATCATCTATCTATGAGGCTCTGCTG.....	1685
517	rgGlyPheValProPheValProGlySerProThrGluArgAsnProPro	533
1686ACCCACAGCGGCGCTCCG	1703
seq_name: gb_in2:AF199463		
seq_documentation_block:		
LOCUS	AF199463	29-NOV-1999
DEFINITION	Drosophila melanogaster sodium-hydrogen exchanger (NHE3), complete cds	
ACCESSION	AF199463	
VERSION	AF199463.1	GI:6470266
KEYWORDS	fruit fly.	
SOURCE	Drosophila melanogaster	
ORGANISM	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
REFERENCE	1 (bases 1 to 2604)	
AUTHORS	Giannakou,M.E. and Dow,J.A.T.	
TITLE	A novel Drosophila member of the NHE exchanger family	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 2604)	
AUTHORS	Giannakou,M.E. and Dow,J.A.T.	
TITLE	Direct Submission	

```
804 ACAGGTACACGATGTGGATGATGAACGAAATCGATCTTTAAGCGAACAGTTTCG 853
      |||
78 erGluAspLeuPhePheIleTyrLeuLeuProPheIleIlePheAsnAla 94
      :::::|||||
854 ATCCGGAGGTATTTTCANATATATTCCTGCGCCCAATATCTTTATATGCG 903
      :::::|||||
95 GlyPheGlnValLysLysGlnPhePheArgAsnPheValThrIleLeu 111
      |||:::|||||
904 GCGTATTCTTGAAGAAAGAAATCTCTTTCGCAACTTGGGTGCCATACT 953
      |||:::|||||
111 tleuPheGlyAlaValGlyThrIleIleSerCysThrIleIleSer.... 126
      |||:::|||||
954 TACGTTTCCCATTTGGGCACGACTTTGTGCGCCTTCCTAATCGGAGGCT 1003
      |||:::|||||
127 .....LeuGlyValThrGlnPhePheLysLysLeuAspIleGlyThrPhe 141
      |||:::|||||
1004 TCATGTACGGCTGTGAACTGATGCCAAGTACTTCAGTAGCAGTTTC 1053
      |||:::|||||
142 AspLeuGlyAspTyrLeuAlaIleGlyAlaIlePheAlaIleThrAspSe 158
      :::::|||||
1054 ACATTCCTGGACTCTTATCTATCTTGGAGCTCTGATATCGCCACAGATCC 1103
      |||:::|||||
158 rValCysThrLeuGlnValLeuAsnGlnAspGluThrProLeu...LeuT 174
      :::::|||||
1104 GCTCACCATTCAGCCATTTCAACGATCTTAGGTCGACGTAAACCTAT 1153
      |||:::|||||
174 yrSerLeuValPheGlyGluGlyValValAsnAspAlaThrSerValVal 190
      |||:::|||||
1154 ATGCGCTAGTCTTGGCGAATCTGTCTCAAGGATGCGGTGCCCATAGTC 1203
      |||:::|||||
191 ValPheAsnAlaIleGlnSerPheAspLeuThrHisLeuAsnHisGluAl 207
      :::::|||||
1204 CTTAGTGGAGCCATTCAAAATTTATGCT...GAACACTATTCCAAATACCGG 1250
      |||:::|||||
207 aAlaPheHisLeu.....LeuGlyAsnPheLeuTyrL 218
      |||:::|||||
1251 GGAATTCGAAATACAGCTTTCTACGCTCGTTGAGCGATTTCTTCTCCA 1300
      |||:::|||||
218 euPheLeuLeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIleSerAla 234
      :::::|||||
1301 TCTTCTTACTATCCCTGATGATGTTGGCGGCCATGGGATGCTTGACAGCA 1350
      |||:::|||||
235 TyrValIleLysLysLeuTyrPheGlyArgHisSerThrAspArgGluVa 251
      :::::|||||
1351 ...TTGATGACCAAAATTTACTCGGGTTCGAGACTTTTCTCTTAGAATC 1397
      |||:::|||||
251 lAlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuP 268
      |||:::|||||
1398 CGGCTCTTCGTGCTCATGAGCTACAGCACTTCTCTGCTGGCTGAGGCTA 1447
      |||:::|||||
268 heAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMetSer 284
      :::::|||||
1448 CGGAACCTACTGGTGTGGCGCTGCTCTCTGCGGCATCTGCCAGGCT 1497
      |||:::|||||
285 HisTyrThrTrpHisAsnValThrGluSerSerArgIleThrThrIlyshi 301
      |||:::|||||
1498 CACTACACGTACAACTTGTGCGAGGACTCGCGCCAAAGGACAAACA 1547
      |||:::|||||
301 sThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrV 318
      |||:::|||||
1548 GATATTCGAGCTCTGAACCTTCTACCGAGAATTTATCTTTTCTCTACA 1597
      |||:::|||||
318 alGlyMetAspAlaLeuAspIleAspLysTrpArgSerValSerAspThr 334
      :::::|||||
1598 TTGGCGTA.....TCGATGTTCACTTT 1620
      |||:::|||||
335 ProGlyThrSerIleAlaValSerSerIleLeuMetGlyLeuVal..... 349
      |||:::|||||
1621 CCCAAGCATCATTTTGACCGCGTTTCATTTATAACAGCTTTTATTTGCGC 1670
      |||:::|||||
350 .MetValGlyArgAlaAlaPheValPheProLeuSerPheLeuSerAsnL 366
      :::::|||||
```


Mon Mar 26 11:04:16 2001

us-09-271-584a-2.p2n.rng

```

1466 .....CACATCTCTCGCCATCGAGGACATATCCGGA.... 1498
491 euThrArgProThrArgThrValHisTyr.....TtpArg 503
1499 .....CAGATCGGCGACAAATATCTCAGAGACAAGTGTCC 1534
504 GlnPheAspSerPheMetArgProVal 513
1535 CACTTCGACAGAGTCTCTCAGCAGGTC 1564
seq_name: /SID56/gcdata/geneseq/NA1993.DAT.Q51524
seq_documentation_block:
ID Q51524 standard; DNA; 4151 BP.
XX
AC Q51524;
XX
DT 18-MAY-1994 (first entry)
XX
DE Rabbit sodium ion/hydrogen ion anti-transport carrier gene.
XX
KW Rabbit; sodium; hydrogen; anti-transport carrier; transgenic mouse;
XX
KW model animal; hypertension; hypercardia; ss.
XX
OS Oryctolagus cuniculus.
XX
FH Key Location/Qualifiers
FT CDS 361..2811
FT FT /*tag= a
XX
PN JP05268856-A.
XX
PD 19-OCT-1993.
XX
PF 26-MAR-1992; 92JP-0068388.
XX
PR 26-MAR-1992; 92JP-0068388.
XX
PA (SUMU) SUMITOMO SEIYAKU KK.
XX
DR WPI: 1993-364207/46.
DR P-PSDB; R44143.
XX
PT Transgenic mouse expressing sodium ion-hydrogen ion
PT anti-transport carrier gene - useful for model animal of
PT hypertension and hypercardia
XX
PS Disclosure: Fig 8-12; 12pp; Japanese.
XX
CC Transgenic mice having this gene are useful as animal model for the
CC analysis of the effect of sodium ion/hydrogen-ion anti-transport
CC carrier gene activation on the regulation of blood pressure. The
CC transgenic animals are useful as models of hypertension and
CC hypercardia.
XX
SQ Sequence 4151 BP; 754 A; 1429 C; 1174 G; 794 T; 0 other;

alignment_scores:
Quality: 441.00 Length: 577
Ratio: 1.353 Gaps: 21
Percent Similarity: 56.499 Percent Identity: 25.303

alignment_block:
US-09-271-584A-2 x Q51524
Align seg 1/1 to: Q51524 from: 1 to: 4151
23 LeuAsnLeuPheValAlaLeuLeuCysAlaCysIleValLeuGlyHisLe 39
673 ATCTCCCTCTGGATCCTGCTGGCTGCCTCATGAAGATAGTTTCCACGT 722

```

```

723 GATCCCAACCATCTCCAGCATCGTCCGGAGAGCTGCCTCTGCTGTTGG 772
55 leGlyLeuGlyThrGlyValThrIleLeuLeuIleSerLysGlyLysSer 71
773 TGGGCTGCTGGTGGGGGCTGATC.....AAGGCGTGGGC 810
72 SerHisLeuLeuValPheSerGluAspLeuPhePheIleTyrLeuLeuPr 88
811 GAGAAAGCGCCCTTCCCTGCGAGTGGAGGTCTTTTCTCTCTTCCTGCTGCC 860
88 opIroIleIlePheAsnAlaGlyPheGlnValLysLysLysGlnPhePheA 105
861 TCCCATCATCTCGAGCGCGCTACTTCTGCGCTGCGCGCAGTTCACCG 910
105 rGAsnPheValThrIleMetLeuPheGlyAlaValGlyThrIle..... 119
911 AGAACCTGGGACCATCTCTATCTTCGCCGTGGTGGCAGCGCTGTGAAC 960
120 .....IleSerCysThrIleIleSerLeuGlyValTh 130
961 GCCTTCTTCCCTGGGCGGCTCATGTACGCCGTGCTGCTGGTGGCGCGGA 1010
130 rGlnPhePheLysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrL 147
1011 GCAG.....ATCAACACATCGGCTCTCTGGACAACC 1042
147 euAlaIleGlyAlaIlePheAlaIleThrAspSerValCysThrLeuGln 163
1043 TGCTGTGGGACAGCATCTCGGCGGTGACCGCGTGGCGTGGCGTGGTGGC 1092
164 ValLeuAsnGlnAspGlu...ThrProLeuLeuTyrSerLeuValPheG 179
1093 GTCTTCGAGGAAATCCACATCAATGAGTGTGTCACATCTCTGGTCTCGG 1142
179 yGluGlyValValAsnAspAlaThrSerValValValPheAsnAlaIleG 196
1143 CGAGTCCCTGCTCAAGCAGCGCTCAGTGTGCTGTATCACCTCTTTG 1192
196 InSerPheAspLeuThrHisLeuAsnHisGluAlaAlaPheHisLeuLeu 212
1193 AGGAGTTT.....GCCAACTACGACCGCTGCGCATCTCTGCTACAGTAC 1236
213 GlyAsnPheLeuTyrLeuPheLeuSer.....ThrLeuLeuG 226
1237 CTGGGCTTCTGAGCTTCTCTGCTGGTGGCGCGGCTGCTGCTGGG 1286
226 yAlaAlaThrGlyLeuIleSerAlaTyrValIleLysLysLeuTyrPheG 243
1287 CGTGGTCTACGGGCTCATCGAGCTTCACTCCCGA.....TTCA 1327
243 lyArgHisSerThrAspArgGluValAlaLeuMetMetLeuMetAlaTyr 259
1328 CCGCCACATCTCCGCGTCAATCGAGCGCTCTCTGCTCTCTACAGTAC 1377
260 LeuSerTyrMetLeuAlaGluLeuPheAspLeuSerGlyIleLeuThrVa 276
1378 ATGGCTTACCTGTCGGCGGAGCTTCTCCACCTGTCGGGCTATCTGGCCT 1427
276 lPhePheCysGlyIleValMetSerHisTyrThrTrpHisAsnValThrG 293
1428 CATCGCTCGGAGTGTGTATGGCGCTTACGTGGAGGCGCAACATCTCC 1477
293 luSerSerArgIleThrThrLysHisThrPheAlaThrLeuSerPheLeu 309
1478 ACAAGTCGACACACCATCAAGTACTTCTCTGAAGATGTGGACGACGTC 1527
310 AlaGluThrPheIlePheLeuTyrValGlyMetAspAlaLeu...AspI 325
1528 AGCAGGACCTCATCTTCTCTCTCTCGGCTCTCCACCGTGGCGGCTC 1577
325 eAsnLysTrpArgSerValSerAspThrProGlyThrSerIleAlaValS 342

```



```

alignment_scores:
  Quality: 415.00
  Ratio: 5.061
  Percent Similarity: 96.471
  Percent Identity: 95.294

alignment_block:
  US-09-271-584A-2 x Z22611 ..

Align seg 1/1 to: Z22611 from: 1 to: 330

358 PheProLeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGluLy 374
      |||||
1 TTTCCGTTATGCTTTCTATCTAACTTAGCCAAAGCAATCAAGCCGAGAA 50
      |||||

374 sileAsnPheasnMetGlnValIleTrpTrpSerGlyLeuMetArg 391
      |||||
51 AATCAACTTTTAACATGCAGGTGTGATTTGGTGGTCTGTCTCATGAGAG 100
      |||||

391 lyAlaValSerMetAlaLeuAlaTyrAsnLysPheThrArgAlaGlyHis 407
      |||||
101 GTGCTGTATCTATGGCTCTTTGCATACAACAAGTTTACAAGGGCGCGAC 150
      |||||

408 ThrAspValArgGlyAsnAlaIleMetIleThrSerThrIleThrValCy 424
      |||||

```

XX
DE Arabidopsis Na/H antiport transporter gene from EST clone T75860.
XX

1156CACTTGATG.....TGCTATTATTTT..... 1176
224 uleuGlyAlaIaThrGlyLeuIleSerAla.....TyrValIleL 238
1177 .CATGGCTGGGATTCACATGATTCCTGCTGTATTACACGCTGCCA 1225
238 ySLyLeuTyrPheGlyArgHis.....SerThrAspArgGluVal 251
1226 GAAGATATATACATACACACTGCTGGATCAGTCTGTATGATACCACTC 1275
252 AlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPh 268
1276 CTCTACATATCCATGGCCCAATTTGCTGCTTTATTTGGTGAATCTTT 1325
268 eAspLeuSerGlyIleLeuThrValPhePheCysGlyIleValMetSerH 285
1326 TTTCCTATTATATTTGACGTTCTTATCACCAGTTAAAGTTACTC 1375
285 lSyrThrTrpHisAsnValThrGluSerSerArgIleThrThrLysHis 301
1376 ACCAAGCAGAACTAATCTGTACATGAAGCT.....GTGAGAGCT 1416
302 ThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrVa 318
1417 ACCTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1461
318 lGlyMetAspAlaLeuAspIleAspLysTrpArgSerValSerAspThrP 335
1462CCATGGGGA.....C 1471
335 roGlyThrSerIleLea.....ValSerSerIleLeu 345
1472 CTCAAGGAAGATTCGAGAAGATATATGATTACATCAATAACATCTC 1521
346 Met.....GlyLeuValMetValGlyArgAlaAlaPheValPhePr 359
1522 ATGCATCATCAGGCTCTCTGGTA.....TCTACAATTTCTGCTTCTT 1565
359 oLeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGlu...LysI 375
1566 TATGAGAGGTTCAAGCAATCTGAGAAGAACTGGAATCAATATAAAA 1615
375 leAsnPheAsnMetGlnValValIleTrpTrpSerGlyLeuMetArgGly 391
1616 TCCAATTTGGA..... 1626
392 AlaValSerMetAlaLeuAlaTyrAsnLysPheThrAlaAlaGlyHisTh 408
1627ACAACTTTTCC.....CACTC 1643
408 rAspValArgGlyAsnAlaIleMetIleThrSerThrIleTr 422
1644 AGATACTCTCGTAGTCATCTACAGGTTTCAACAATCAGC 1686

seq_name: /cn2_6/ptdata/1/ina/5A_COMB.seq:us-08-098-141-1

seq_documentation_block:
; Sequence 1, Application US/08098141
; Patent No. 5441875
; GENERAL INFORMATION:
; APPLICANT: Hediger Ph.D., Matthias A.
; TITLE OF INVENTION: Urea Transporter Polypeptide
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Choate, Hall & Stewart
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02119
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098.141
FILING DATE: 19930723
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kaplan Esq., Warren A.
REGISTRATION NUMBER: 34,199
REFERENCE/DOCKET NUMBER: 092662-001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-227-5020
TELEFAX: 617-227-7566
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3060 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-098-141-1

alignment_scores:
Quality: 98.50 Length: 319
Ratio: 0.724 Gaps: 17
Percent Similarity: 42.633 Percent Identity: 21.317

alignment_block:
US-09-271-584A-2 x US-08-098-141-1
Align seg 1/1 to: US-08-098-141-1 from: 1 to: 3060

34 IleValLeuGlyHisLeuLeuGluGluAsnArgTrpMetAsnGluSerIl 50
1362 ATCTGATGGC...CTCTTTGCCAGAACCCATGGTGG.....GCCAT 1302
50 eThrAlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeuLeuL 67
1303 CGCGGGATGCTGGGACCGCTGATGTCACCTTGACCTGCCCTCATCTGA 1352
67 erLysGlyLysSerSer.....His..... 73
1353 GCCAGGACAGGTACGCCATCCCTCCGACTCCACGGCTACACGGCGTG 1402
74LeuLeuValPheSerGlu.....AspLeuPhePh 83
1403 CTGGTGGGCTGCTGATAGCGGTGTTCTCCGACAAAGGCGGACTATTACTG 1452
83 eIleTyrLeuLeuProPheIleIlePheAsnAlaGlyPheGlnValLysL 100
1453 GTGGCTTCTGCTCCCTGTCATCGTGATGTC..... 1483
100 ySLysGlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaVal 116
1483 1483
117 GlyThrIleIleSerCysThrIleIleSerLeuGlyValThrGlnPhePh 133
1484ATGCTCTCCCATCTCTCCAGTGCCTGGGCGGACCATCTT 1524
133 eLysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleG 150
1525 CAGCAAGTGGGACCTCCAGGCTTCCACACTGCCCTTCAACATCGCCGTGA 1574
150 lYAlaIlePheAlaThrAspSerValCysThrLeuGlnValLeuAsn 166
1575 CCCTGTACTTGGCAGCCACAGGCCAC..... 1600
167 GlnAspGluThrProLeuLeuTyrSerLeuValPheGlyGluValVa 183


```

114 GlyAlaValGlyThrIleSerCysThrile.....IleSe 126
...:|||||:|||||:|||||:|||||:|||||:|||||:
334 GCTATCATCGGAACCTTTCTGCTCAGTTGTTACTGGTGGTCTGGTTA 383
126 rLeuGlyValThrGlnPhePheLeuLysLeuAspIleGlyThrPheAspL 143
|||||:|||||:|||||:|||||:|||||:|||||:
384 TCTTGGCGCTCTATGATCTC.....ATGATAAAGCTCCCT 421
143 euGlyAspTyrLeuAlaIleGlyAlaIlePheAlaAlaThrAspSerVal 159
...:|||||:|||||:|||||:|||||:|||||:|||||:
422 TTGTTGAGTGTCTATGTTGTTGGTCACTATATACGCTACGACCCCTGTC 471
160 CysThrLeuGlnValLeuAsnGlnAspGluThrProLeu...LeuTyrSe 175
...:|||||:|||||:|||||:|||||:|||||:|||||:
472 ACTGTACTCTCTATATTCAGGATGCGGACCGATGTTAACTCTGATGC 521
175 rLeuValPheGlyGluGlyValValAsnAspAlaThrSerValValValp 192
...:|||||:|||||:|||||:|||||:|||||:|||||:
522 TTGTTGCTTTGGGAATCAGTTCTCAATGATGCTATGCAATATCATTTGT 571
192 heAsnAlaIleGlnSerPheAspLeuThrHisLeuAsnHisGluAlaAla 208
...:|||||:|||||:|||||:|||||:|||||:|||||:
572 ACAGAAATGTCCTTAGTAACCCGACGCTCTGCTGGGGAACATTTT 621
209 PheHisLeuLeuGlyAsnPheLeuTyrLeuPheLeuLeuSerThrLeuLe 225
|||||:|||||:|||||:|||||:|||||:|||||:
622 TTCATGGTGGTATCATAGGTTTTCGAGACTTTTTCGCTGCTCAATGCTGTC 671
225 uGlyAlaAlaThrGlyLeuIleSerAlaTyrValIleLysLysLeuTyrp 242
|||||:|||||:|||||:|||||:|||||:|||||:
672 AGGGTTGGGTTGGATTACCTTCAGCTTACTCTTTAAG.....TATG 715
242 heGly.....ArgHisSerThrAspArgGluValAlaLeuMetMet 255
|||||:|||||:|||||:|||||:|||||:|||||:
716 CAGGATGGACACCGAGATCTTCAGAACTTGGAGTGTGCTCTTTGTA 765
256 LeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPheAspLeuSerGl 272
|||||:|||||:|||||:|||||:|||||:|||||:
766 CTTTCCCGTATTTTCATACATGCTTCGACAGAGGTGTGGTCTCTCCGG 815
272 yIleLeuThrValPhePheCysGlyIleValMetSerHisTyrThrTrph 289
|||||:|||||:|||||:|||||:|||||:|||||:
816 CATTTGTTCTATCTCTTCACAGGAATGTTATGAAGCGCTACACTTTCT 865
289 isAsnValThrGluSerSerArgIleThrThrLysHisThrPheAlaThr 305
|||||:|||||:|||||:|||||:|||||:|||||:
866 CAATCTCTCAGAGCTTCACAGAGTTTCGTATCTCTTTTTCCTTG 915
306 LeuSerPheLeuAlaGluThrPheIlePheLeuTyrValGlyMetAsp.. 321
...:|||||:|||||:|||||:|||||:|||||:|||||:
916 ATATCTTCGTACAGAACTTTCACGTTCACTTATACATGGGATTTGATAT 965
322 .AlaLeuAspIleAspLysTyrArgSerValSerAspThrProGlyThrS 338
|||||:|||||:|||||:|||||:|||||:|||||:
966 TGGCATGGAGCAGTAGCTGGTCCCAT..... 993
338 erIleAlaValSerSerIleLeuMetGlyLeuVal..... 349
|||||:|||||:|||||:|||||:|||||:|||||:
994 .....GTGGGTTATCTCTTCTCTATGTTATCTCTATCTCTATCTGAT 1035
350 .....MetValGlyArgAlaAlaPheValPheProLeuSerPheLe 363
...:|||||:|||||:|||||:|||||:|||||:|||||:
1036 CGTCATGTATGTATGTCAGTGGCTGCTCAATGTTATTTGGGTGTCATATTT 1085
363 userAsnLeuAlaLysAsnGlnSerGluLysIleAsnPheAsnMetG 380
| |||||:|||||:|||||:|||||:|||||:|||||:
1086 GGTCAACCTATTAGACAG...GAGAACGAGAGATACCTATGAGCACCC 1132
380 nValValIleThrTrpSerGlyLeuMetArgGlyAlaValSerMetAla 396
||:|||||:|||||:|||||:|||||:|||||:|||||:
1133 AAAAGCCCTTTGGTATAGTGGACTT...CGAGGGCAATGGCATTTGCA 1179

```

```

397 LeuAla.TyrAsnLysPheThrArgAlaGlyHisThrAspValArgGlyA 413
|||||:|||||:|||||:|||||:|||||:|||||:
1180 CTTGCACCTTCATCATCTATCATGATCATCAGAGGGTCACGCCAAATCAT 1229
413 snAlaIleMetIleThrSerThrIleThrValCysLeuPheSerThrVal 429
...:|||||:|||||:|||||:|||||:|||||:|||||:
1230 CT.....TTACTGCAACCACTATTTGTTGTTCACGGTTTA 1270
430 ValPheGlyMetLeuThrLysProLeuIleSerTyrLeuLeuProHisGl 446
...:|||||:|||||:|||||:|||||:|||||:|||||:
1271 CTAATAGGAGGTCGACAGGTAATAATCTTGAAGCTTTG..... 1309
446 nAsnAlaThrThrSerMetLeuSerAspAspAsnThrProLysSerIleH 463
...:|||||:|||||:|||||:|||||:|||||:|||||:
1310 .....GAAAGTTGAGTGACGAT..... 1327
463 isIleProLeuLeuAspGlnAspSerPhe.....IleGluProSer 476
|||||:|||||:|||||:|||||:|||||:|||||:
1328 .....CTTGAT...GACTCCATGCTGAAGGCTTTGAAGAGAGC 1363
477 GlyAsnHisAsnValProArgProAspSerIleArgGlyPheLeuThrAr 493
...:|||||:|||||:|||||:|||||:|||||:|||||:
1364 GATCATCATGATGTCCTCTCTCTTTAGCATTTGGAGCTTCATCTGACGA 1413
493 gProThr..... 499
|||||:|||||:|||||:|||||:|||||:|||||:
1414 GGATACATCATCATCAGAGAGGTTCAAGATCAAGCTGAAGAGTTTC 1463
499 isTyrTyrTrpArgGlnPhe.....AspAspSerPheMetArgPro 512
|||||:|||||:|||||:|||||:|||||:|||||:
1464 ACAAAACCACTACATCATTCACCGCTGGACAAAAAACTTCTGACTCCG 1513
513 ValPhe 514
|||||
1514 TTCCTC 1519

seq_name: /SIDS6/gcgdata/geneseq/geneseqn/NA1997.DAT:T51330
seq_documentation_block:
ID T51330 standard; CDNA; 2584 BP.
XX
AC T51330;
XX
DE 29-APR-1997 (first entry)
XX
DE Human Na+/H+ exchanger isoform NHE3 composite cDNA.
XX
KW NHE3; Na+/H+ exchanger; sodium ion-proton exchanger; antiporter;
KW chronic metabolic acidosis; chronic renal failure; hypertension;
KW diabetic nephropathy; familial diarrhoea; cystic fibrosis; kidney;
KW small intestine; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 11..2515
FT /tag= a
FT /transl_except= pos:1970..1972..aa:Arg
FT /note= "codon CAC (bases 1870-1972) codes for Arg"
XX
PN WO9703196-A1.
XX
PD 30-JAN-1997.
XX
PF 10-JUL-1996; 96WO-US11670.
XX
PR 11-JUL-1995; 95US-0001061.
XX
PA (MAGA-) MAGAININ PHARM INC.
XX
PI Brant SR, Donowitz M, Tse C, Yun CHC;
XX
DR WPI; 1997-119051/11.

```

```
1156 .....CACTGATG.....TGGTATTATTTT..... 1176
224 uNeuGlyAlaAlaThrGlyLeuSerAla.....TyrValIleL 238
1177 ..CACTGGCTGGGATTTCCATTGATTCCTGTTATTCACGCTGTGCCA 1225
238 ySLyLeuTyrPheGlyArgHis.....SerThrAspArgGluVal 251
1226 GAAGATTATATCAATGACAACTGCTGATCAGTCTGTGATACCAACTC 1275
252 AlaLeuMetMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPh 268
1276 CTCACATATCCATGCGCCCAATTTGCTGCTTTTATTGTTGAATCTTT 1325
268 eAspLeuSeuGlyIleLeuThrValPhePheCysGlyIleValMetSerH 285
1326 TTTCTATTATATTTGATGCTGCTGCTTATACCAAGTTAAAGTTACTC 1375
285 IsTyrThrTrpHisAsnValThrGluSerSerArgIleThrThrLysHis 301
1376 ACCAAGCAGAACTCAATCTGTATCAAGAACT.....GTGAGAGCT 1416
302 ThrPheAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrVa 318
1417 ACGTTATCTCTGGTGGCTGCTGCTTGGCATTTGAATTTGCTGATT..... 1461
318 lGlyMetAspAlaLeuAspLeuAspLysTrpArgSerValSerAspThrp 335
1462 .....CCATGGCGA.....C 1471
335 roGlyThrSerIleAla.....ValSerSerIleLeu 345
1472 CTGAAGAGAGATTCAGAGAGATATATGATTCATCAATAACATCCTC 1521
346 Met.....GlyLeuValMetValGlyArgAlaAlaPheValPhePr 359
1522 ATGCACTATCAGGGTCTCTTGTA.....TCTACAATTTTCTGCTCTT 1565
359 oLeuSerPheLeuSerAsnLeuAlaLysLysAsnGlnSerGlu...LysI 375
1566 TAAAGGAGAGGTTCAAGCAATCTCGAAGAACTGGGAATCAATATAA 1615
375 leAsnPheAsnMetGlnValValIleTrpTrpSerGlyLeuMetArgGly 391
1616 TCCAATTTGGA.....AACAACTTTTCC.....CACTC 1643
392 AlaValSerMetAlaLeuAlaTyrAsnLysPheThrArgAlaGlyHisTh 408
1627 ..... 1626
408 AspValArgGlyAsnAlaIleMetIleThrSerThrIleThr 422
1644 AGATACCTCCGTAGTGCATTCACACGGTTTCAACATCAGC 1686
```

seq_name: /cgn2_6/ptodata/1/ina/5A_COMB.seq:US-08-098-141-1

seq_documentation_block:

Sequence 1, Application US/08098141

Patent No. 5441875

GENERAL INFORMATION:

APPLICANT: Hediger Ph.D., Matthias A.

TITLE OF INVENTION: Urea Transporter Polypeptide

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSER: Choate, Hall & Stewart

STREET: Exchange Place, 53 State Street

CITY: Boston

STATE: MA

COUNTRY: U.S.A.

ZIP: 02119

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/098,141
FILING DATE: 19930723
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kaplan Esq., Warren A.
REGISTRATION NUMBER: 34,199
REFERENCE/DOCKET NUMBER: 092662-001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617:227-5020
TELEFAX: 617:227-7566
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3060 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-098-141-1

alignment_scores:
Quality: 98.50 Length: 319
Ratio: 0.724 Gaps: 17
Percent Similarity: 42.633 Percent Identity: 21.317
```

alignment_block:

US-09-271-584A-2 x US-08-098-141-1

Align seg 1/1 to: US-08-098-141-1 from: 1 to: 3060

```
34 IleValLeuGlyHisLeuLeuGluAsnArgTrpMetAsnGluSerI 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1262 ATCGTGATTGGC...CTCTTTGTCAGAACCCATGTTGG...GCCAT 1302
50 ePheAlaLeuLeuIleGlyLeuGlyThrGlyValThrIleLeuLeu 67
|||||:|||||:|||||:|||||:|||||:|||||:
1303 CGCGGATGCTGGCACCGTCATGTCACCTGACCTGCCTCATCTGA 1352
67 eLysGlyLysSerSer.....His..... 73
|||||:|||||:|||||:|||||:|||||:
1353 GCCAGCACAGGTACGATCGCTCCGAGCTCCACGGCTACAACGGCGTG 1402
74 .....LeuLeuValPheSerGlu.....AspLeuPhePhe 83
|||||:|||||:|||||:|||||:|||||:
1403 CTGGTGGGCTGCTGATAGCGGTCTTCCGACAAAGGCGGCTATTACTG 1452
83 eLeuTyrLeuLeuProPheIlePheAsnAlaGlyPheGlnValLysL 100
|||||:|||||:|||||:|||||:|||||:
1453 GTGGCTTCTGCTCCCTGTCATCGTGTGATGCC..... 1483
100 ySLysGlnPhePheArgAsnPheValThrIleMetLeuPheGlyAlaVal 116
1483 ..... 1483
117 GlyThrIleIleSerCysThrIleIleSerLeuGlyValThrGlnPhePhe 133
|||||:|||||:|||||:|||||:|||||:
1484 .....ATGTCCTGCCCATCTCTCCAGTGCCCTGGGACCATCTT 1524
133 eLysLysLeuAspIleGlyThrPheAspLeuGlyAspTyrLeuAlaIleG 150
|||||:|||||:|||||:|||||:|||||:
1525 CAGCAAGTGGGACCTCCCGAGTCTTCCACTGCCCTTCAACATCGCGTGA 1574
150 lyAlaIlePheAlaThrAspSerValCysThrLeuGlnValLeuAsn 166
|||||:|||||:|||||:|||||:
1575 CCCTGTACTCTGGCAGCCACAGGCCAC..... 1600
167 GlnAspGluThrProLeuLeuTyrSerLeuValPheGlyGluGlyValVa 183
```



```
1601 .....TACACCTCTTCCACACACTGCT 1629
183 lAsnAspAlaThrSerVal.....ValValPheAsnAlaIleGlnSerP 198
1630 GCAGCGCTGATCCTCCGTCACACATCAGCTGGTCAGAGATCCAG.... 1675
198 heAspLeuThrHisLeuAsnHisGluAlaAlaPheHisLeu..... 212
1676 .....GTGCCCTGCTCTTGAGA 1693
212 ..... 212
1694 GCCATCCCGTGGGAATGGCCAGGTATGGCTGTGATAACCCCTGGAC 1743
213 ....GlyAsnPheLeuTyr...LeuPheLeuLeuSerThrLeu..... 224
1744 TGGAGGCATTTTCTCCTGCTGCTGTATATCTTCACTCTTATTGCT 1793
225 .....LeuGlyAlaAlaThrGlyLeuIleSerAlaThrValIle 237
1794 TGCATGCGGCAATGGATGCCATGCGGATGTAGCAGCATCAGGATT 1843
238 .....LysLeuLeuTyrPheGlyArgHisSerThrAspArgG 250
1844 GCCACACCCCTTTGACTCCATCTATTGGCTGTGTGGCTTCAACAGCAC 1893
250 uValAla.....LeuMetMetLeuMetAlaThrLeuS 261
1894 GCTGCATGATCCCTGCTGGAGGATGTTCTACGTATCAGCTGGCAGA 1943
261 erTyrMetLeuAla.....GluLeuPhe..... 268
1944 CTCACCTCTTCTGCTGCTGGCTGGCTGCTGCTGGCTGCTGCTGCT 1993
269 AspLeuSerGlyIleLeuThrValPhe..... 277
1994 GCCCTGACCAAGTATTATCTGTGTTGGATTACCAACCTGCACCTGCC 2043
278 PheCys 279
2044 CTTCTGC 2050
```

seq_name: /cgn2_5/ptodata/1/ina/6_COMB.seq:US-08-602-809-3

seq_documentation_block:

; Sequence 3, Application US/08602809

; Patent No. 6008012

; GENERAL INFORMATION:

; APPLICANT: BERGSMAN, DEREK

; APPLICANT: ELLIS, CATHERINE

; TITLE OF INVENTION: HUMAN SOMATOSTATIN-LIKE R

; TITLE OF INVENTION: RECEPTOR

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESS: Ratner & Prestia

; STREET: P.O. Box 980

; CITY: Valley Forge

; STATE: PA

; COUNTRY: US

; ZIP: 19482-0980

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/602,809

; FILING DATE: 13-JUN-1997

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/16472

; FILING DATE: 15-DEC-1995

; APPLICATION NUMBER: US 08/357,675
; FILING DATE: 16-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestia, Paul F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: P50277
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 601-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1080 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-602-809-3

alignment_scores:

Quality: 96.50 Length: 371

Ratio: 0.652 Gaps: 22

Percent Similarity: 39.892 Percent Identity: 22.642

alignment_block:

US-09-271-584A-2 x US-08-602-809-3

Align seg 1/1 to: US-08-602-809-3 from: 1 to: 1080

```
171 ProLeuLeuTyrSerLeuValPheGlyGluGlyValValAsnAspAlaTh 187
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
94 CCTACCTCTACAGCATCTCTTGTGTGGGGAATA..... 129
187 rSerValValPheAsnAlaIleGlnSerPheAspLeuThrHisLeuA 204
129 ..... 129
204 snHisGluAlaAlaPheHisLeuLeuGlyAsnPheLeuTyrLeuPheLeu 220
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
130 .....TTTGGAAACACAGCTTGGTGGTCACTTC 156
221 LeuSerThrLeuLeuGlyAlaAlaThrGlyLeuIleSerAlaTyrValI 237
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
157 ATT.....TACTTTTA 167
237 eLysLysLeuTyrPheGlyArgHisSerThrAspArgGluValAlaLeuM 254
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
168 CATGAAGCTG.....AGACTGTGGCCAGCG 193
254 etMetLeuMetAlaTyrLeuSerTyrMetLeuAlaGluLeuPheAspLeu 270
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
194 TCTTTCTT.....CTCAATCTCGCTTGGCT.....GACTTGA 225
271 SerGlyIleLeuThrValPhePheCysGlyIleValMetSer...His 286
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
226 TGCCTTTTCTGACTTTCCTCTGCGGAGCTATATACCCCTATGAGTA 275
286 rThrTrpHisAsnValThrGluSerSerArgIleThrThrLysHisThrP 303
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
276 CGCTGGCCCTTCGGCAATCACCCTATATGATGATCGCTTCG..... 315
303 heAlaThrLeuSerPheLeuAlaGluThrPheIlePheLeuTyrValGly 319
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
316 ..GCCAGCTGAGCTTCAACCTCTACCGCAGCTGTGTCTCTT..... 354
320 MetAspAlaLeuAspIleAspLysTrpArgSerValSerAspThrProG 336
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
355 CTCACGTGTCTCAGCATCAGCCGCTAC..... 381
336 yThrSerIleAlaValSerSerIleLeuMetGlyLeuValMetValGlyA 353
381 ..... 381
```

